

The Neurobiology of Neural Networks

Computational Neuroscience

Terrence J. Sejnowski and Tomaso A. Poggio, editors

Methods in Neuronal Modeling: From Synapses to Networks, edited by Christof Koch and Idan Segev, 1989

Neural Nets in Electric Fish, Walter Heiligenberg, 1991

The Computational Brain, Patricia S. Churchland and Terrence J. Sejnowski, 1992

Dynamic Biological Networks: The Stomatogastric Nervous System, edited by Ronald M. Harris-Warrick, Eve Marder, Allen I. Selverston, and Maurice Moulins, 1992

The Neurobiology of Neural Networks, edited by Daniel Gardner, 1993

The Neurobiology of Neural Networks

edited by
Daniel Gardner

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Copyrighted Material

© 1993 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set in Palatino by Daniel Gardner and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

The Neurobiology of neural networks / edited by Daniel Gardner.

p. cm. — (Computational neuroscience)

"A Bradford book."

Includes bibliographical references and index.

ISBN 0-262-07150-9

1. Neural circuitry. 2. Neural networks (Computer science)

3. Neurobiology. I. Gardner, Daniel, 1945- II. Series.

QP363.3.N455 1993

612.8—dc20

93-9125

CIP

Contents

Contributors	ix
Preface	xi
1. Introduction: Toward Neural Neural Networks <i>Daniel Gardner</i>	1
2. Two Principles of Brain Organization: A Challenge for Artificial Neural Networks <i>Charles F. Stevens</i>	13
3. Static Determinants of Synaptic Strength <i>Daniel Gardner</i>	21
4. Learning Rules from Neurobiology <i>Douglas A. Baxter and John H. Byrne</i>	71
5. Realistic Network Models of Distributed Processing in the Leech <i>Shawn R. Lockery and Terrence J. Sejnowski</i>	107
6. Neural and Peripheral Dynamics as Determinants of Patterned Motor Behavior <i>Hillel J. Chiel and Randall D. Beer</i>	137
7. Dynamic Neural Network Models of Sensorimotor Behavior <i>Eberhard E. Fetz</i>	165
References	191
Index	219

Series Foreword

Computational neuroscience is an approach to understanding the information content of neural signals by modeling the nervous system at many different structural scales, including the biophysical, the circuit, and the systems levels. Computer simulations of neurons and neural networks are complementary to traditional techniques in neuroscience. This book series welcomes contributions that link theoretical studies with experimental approaches to understanding information processing in the nervous system. Areas and topics of particular interest include biophysical mechanisms for computation in neurons, computer simulations of neural circuits, models of learning, representations of sensory information in neural networks, systems models of sensory-motor integration, and computational analysis of problems in biological sensing, motor control, and perception.

Terrence J. Sejnowski
Tomaso Poggio

Contributors

Douglas A. Baxter
Department of Neurobiology and Anatomy
The University of Texas Medical School
Houston, Texas

John H. Byrne
Department of Neurobiology and Anatomy
The University of Texas Medical School
Houston, Texas

Hillel J. Chiel
Department of Biology
and Department of Neuroscience
Case Western Reserve University
Cleveland, Ohio

Randall D. Beer
Department of Computer Engineering and Science
and Department of Biology
Case Western Reserve University
Cleveland, Ohio

Eberhard E. Fetz
Department of Physiology and Biophysics
and Regional Primate Research Center
University of Washington
Seattle, Washington

Daniel Gardner
Department of Physiology
and Department of Neuroscience
Cornell University Medical College
New York, New York

Shawn R. Lockery
Computational Neurobiology Laboratory
The Salk Institute for Biological Studies
San Diego, California

Terrence J. Sejnowski
Computational Neurobiology Laboratory
The Salk Institute for Biological Studies
San Diego, California

Charles F. Stevens
Howard Hughes Medical Institute
The Salk Institute for Biological Studies
San Diego, California

Preface

The Neurobiology of Neural Networks is intended not only as a reference for students and investigators, but as a spur and a guide toward development of the next generation—the third generation—of neural network models. It is a series of variations on a single theme: that both artificial neural networks and neurobiology will benefit by a partial synthesis of their disparate natures. Individual chapters were commissioned from selected authors with this goal in mind; this work is *not* the proceedings of a specific meeting or an ad hoc collection of unrelated works. Each chapter examines neural network assumptions and models from a neurobiological perspective, presenting a critical review of the experimental evidence both consistent and inconsistent with neural network components, mechanism, architecture, and capabilities. Chapters are designed and written so that readers who are familiar with neurobiology but not network models, or the reverse, can appreciate the synthesis. To enhance readability, each chapter is self-contained, resulting in parallel or complementary views of similar topics. Frequent cross-references between chapters emphasize the unity of the theme, and a single combined bibliography emphasizes the unity of purpose.

Neural networks need input, and the field of neural networks needs continuing input from both of the two groups for whom this volume is intended: neurobiologists and network modelers.

For neurobiologists, the book presents a self-contained analysis of the methodology, strengths and weaknesses, and significance of artificial neural network models and their applicability to neuroscience. Because no single volume can be sufficient, there are plentiful references to the neural network literature, concentrating on the computationally powerful but inadequately neuromorphic nets developed within the last decade. Throughout, the perspective is neurobiological; all chapters have one or more authors who are practicing as well as professing neurobiologists, using single-unit or intracellular recording, voltage- or patch-clamping, neurochemical or molecular biological techniques. A major goal is to encourage neurobiologists to participate in constructing the next generation of neural networks, as well as to examine neurobiological data from a network perspective.

For neural network modelers, especially those without a background in the life sciences, the book critically examines aspects of network models that are likely to be both computationally significant and comparable to our current view of real neurons and synapses and the networks they form. Because no single volume can be sufficient, there are plentiful references to the relevant literature in neurobiology, concentrating on significant developments in neuronal biophysics, synaptic plasticity, and interneuronal organization within the last decade. Although the perspective is neurobiological, all chapters have one or more authors who have come to neurobiology from the physical sciences or engineering, or who have broadened their perspective by a detour into these fields. A major goal is to encourage neural network modelers to participate in constructing the next generation of neural networks, as well as to examine network design and performance from a more neurobiological perspective.

Specific acknowledgements follow individual chapters, and include a large number of diverse public and private funding agencies: the Office of Naval Research, the National Institutes of Health, the National Institute of Mental Health, the Air Force Office of Scientific Research, the National Science Foundation, the Howard Hughes Medical Institute, The Bank of America-Giannini Foundation for Medical Research, The Mathers Foundation, and the Cleveland Advanced Manufacturing Program. Although each agency has a distinct mission, the support each has provided to this emerging field is a tribute not only to the universality of these investigations, but also to the vision of many peer reviewers and program officers.

One final acknowledgement is general to the book, yet specific to its editor: I thank Deborah and Ben, and especially Esther, for their support during the many iterations and recursions needed to tune this work to the desired output pattern.

Daniel Gardner
New York City
Thanksgiving Day, 1992

The Neurobiology of Neural Networks

1

Introduction: Toward *Neural* Neural Networks

Daniel Gardner

NEURON AND NETWORK

It is a truth universally acknowledged, that a single neuron in possession of a good potential must be in want of a network. Universally acknowledged as well is an ever-increasing list of the properties of single neurons, but neurons perform no useful task in isolation. We are in want of an understanding of the properties of neurons in networks, how each neuron contributes to a neuronal array and what properties emerge as a result of the interconnection, and how the network performs the information processing tasks that underlie the function of the nervous system and the behavior of the organism.

The need for understanding network properties can be appreciated by considering a global view of the nervous system and how it differs from other organ systems. Compared to other organ systems, the nervous system is unique. It transcends the purely mechanistic functions of other systems of the body. Other systems pump fluids, exchange gases, or concentrate urine. The nervous system observes, reasons, composes, reflects, ponders, deduces, hesitates, and acts. Yet the nervous system produces these results by utilizing physiological and biophysical mechanisms that resemble those used in the other organ systems. Understanding this interplay between mechanism and effects requires understanding not only how the cellular and subcellular components of the nervous system operate, but also how they are put together. There is increasing evidence that it is the network and connective properties of small sets of neurons which form the basis for nervous system function.

TOWARD A NEW GENERATION OF NEURAL NETWORKS

The First and Second Generations

Artificial neural network models are currently in their second generation. These networks were composed of interconnected arrays of individual processing units, called elements or neurons, that summed and transformed input signals to produce binary or graded signals of output

activity. Appropriate adjustment of variable connection coefficients, or *synaptic weights*, transformed input patterns—encoded as activity of several units—into corresponding output patterns. The first generation, typified by the perceptron (Minsky and Papert 1969; Rosenblatt 1958) was loosely based on early concepts of biological neurons that antedate much contemporary knowledge of neuronal and synaptic function. Both elements and architecture of these first-generation models thus lacked *neuromorphism*. They also lacked significant capabilities; the classic example of Minsky and Papert (1969) was an inability to encode the basic exclusive-or operation of elementary logic.

The second generation differs architecturally, technologically, and functionally, but not neuromorphically. Much of the second generation is characterized by a multilayer architecture, a connectivity separating input, intermediate, and output units. These networks are trained by repetitive presentation with a set of input data in order to tune synaptic weights to appropriate values. Following training, the now-appropriate weights are used to process input information. This generation is made possible by several recent successful solutions to the problem of *credit assignment*. This problem (actually one of blame assessment) asks how the error signals characterizing poorly functioning networks are to be distributed to individual synapses and so minimized. The backpropagation algorithm of Rumelhart et al. (1986a) is a well-known and largely successful solution to the credit assignment problem and is discussed throughout this volume; a possible neurobiological analog of this solution is examined in chapter 3.

Technology has aided interest in second-generation networks and their application; the prevalence of inexpensive and powerful desk-top computers has empowered a host of network modelers. Most important, these networks have proven functional, capable of solving impressive problems in information storage, processing, and transformation. In spite of these increased capabilities, they remain restricted to certain classes of problems, they are difficult to train and to scale, and they remain as non-neuromorphic as the first generation.

Third Generation Neural Networks Should Be Neuromorphic

Is a network of neurons a neural network? This book calls for a new generation of neural networks. Networks of biological neurons, comprising the nervous system of a behaving organism, transcend the simplified components, layered architecture, limited scale, and restricted capabilities of first- and second-generation artificial network models. A goal of this volume is to explore paths by which artificial network models can be made more neuromorphic, and thereby advance these models so that it will not require oversimplification or wishful thinking to say that a network of neurons is a neural network.

Understanding and incorporating relevant neuromorphism into

models requires reverse-engineering the nervous system on each of several levels. The analysis of neuronal and subneuronal components itself proceeds on several levels, including the molecular, channel, membrane, cellular, and synaptic. Selective addition of neuromorphic features from these studies presents a continuing challenge to development of third-generation networks. However, neural network models consist of more than their elements, and third-generation networks should employ neuromorphic connectivity as well.

Elements and connectivity are related in nervous systems, and successful third-generation networks are likely to exploit these relations. It may no longer be possible to view the process of model building, or even of reverse-engineering the nervous system, as capable of being compartmentalized into such discrete levels as molecular, dendritic, cellular, network, columnar, and regional. Such analysis is possible when the state variables characterizing one level are understood and their lower-level implementation, or higher-level exploitation, is to be discovered. The nervous system is unlikely to yield to this restricted discrete approach. Synapses do more than transfer discrete amounts of current into a cell. Neuronal state depends on more than the binary presence or absence of an action potential at fixed time points, or even more than an average rate of firing. Network function and even effective connectivity can be modulated by external and internal factors, including specific molecules and neurons. Simplified encapsulation of the most apparent factors of each of these levels may permit easier modeling of the next higher array, but such models are likely to be incomplete, insufficient, or irrelevant.

PARENTS OF THE THIRD GENERATION

The third generation, then, will be a joint product of neurobiology and neural networks. Like many contemporary couples, neurobiology and neural networks seem to thrive without offspring. As a field of endeavor, neurobiology is matched only by molecular biology in the rapid increase in fundamental understanding. Similarly, advances in design, efficiency, computability, and application characterize the field of neural networks. What are the advantages to each parental field of a hybrid?

The Third Generation's Potential for Neurobiology

In spite of a continuing outpouring of significant and exciting findings, we still do not understand how the brain works. Two lines of evidence suggest the importance of future-generation network models in aiding this understanding. Distributed processing algorithms characterize much of what is known about nervous system function, but contemporary methods are not well suited for acquiring data for testing and interpreting these algorithms. Single-unit recording, even in its most technically

sophisticated patch-clamp variants, can provide information only about individual neurons. Evoked-potential studies, even microelectrode field-potential recordings, indiscriminately combine signals from larger or smaller classes of cells. New data acquisition methods such as optical imaging (Morton et al. 1991) or multiple-electrode arrays (Mountcastle et al. 1992) are capable of recording simultaneously from many elements of a network, but understanding data acquired by these methods will require new metaphors, new paradigms—in short, new models (Gerstein and Turner 1990).

Functional interneuronal organization is more likely than other aspects of the nervous system to be aided by the multidisciplinary contributions of third-generation models. Such advances would not be expected, for example, in the biophysics of neuronal components. Molecular components of neurons are restricted by biophysics and by protein biochemistry (Hille 1992). It would be unrewarding to model single Ca^{2+} -selective membrane channels with continuously variable conductances, because such channels are structurally implausible. However, network wiring schemes and architecture are less restricted by basic physically derived principles, deriving their organization more from opportunistic and less-constrained evolutionary divergence.

Appropriate network analogs with organizational free parameters, exploring their parameter space much as evolution may have explored its, may therefore prove consistent with, even predictive of, neurobiology. The freedom of 'experimental' manipulation permitted by future models can supplement the more restricted alterations that neuroscience allows. In spite of such recent advances as the introduction of modifiable links to a biological network by the use of a variable artificial conductance (Sharp et al. 1992), neurobiological systems more readily permit examination of connectivity than its modification. Single cells can be stimulated or suppressed with intracellular stimulation, lesions ranging from micro to macro can be produced (see chapters 6 and 7), and neuromodulators introduced, but these manipulations are usually restricted in scope, targets, precision, and repeatability. Synaptic plasticity can be induced or exploited, but only to a limited degree. Neurons can be grown in culture to produce networks of indiscriminate or partially defined connectivity, but cellular and synaptic as well as network properties may differ from those in vivo. Network model manipulations have no such restrictions, and this freedom itself suggests the importance of neuromorphism in selecting both parameters to be varied, and the appropriate ranges.

The Third Generation's Potential for Neural Network Design

Neuromorphic third-generation networks are likely to provide significant advances in processing power, and this should be of obvious interest to those who wish to exploit the capabilities of these nets, even

without a direct interest in nervous system function. Limitations of present-generation networks are well known; some are listed above. Similarly well known is the processing power of biological networks, including the human brain: a massively parallel processor implementing multiple-instruction, multiple-data-stream protocols. One central tenet of third-generation networks is that incorporating more of the principles and algorithms of biological neural networks may increase the capabilities of their artificial analogs.

Independent of their lack of similarity to networks of real neurons, neural network models have computational ability. Moreover, these models are internally consistent within their own domain, and it may seem unfair to ask that they be constrained (or expanded) by an additional set of mathematically arbitrary but biologically consistent parameters. Nevertheless, one of the goals of this volume is the modification of existing neural network models in the hope that increases in neuromorphic complexity will lead to parallel advances in power.

If present-generation network models can perform as well as they do using oversimplified units and synapses, how much more will be possible when they incorporate increasing amounts of the rich and complex repertoire of neuronal mechanisms? Neuromorphism can be viewed as a series of evolutionary compromises needed to permit cellular mechanisms to function at some small fraction of the speed, repeatability, and precision of a logic gate. I reject this view, and believe that most neurobiologists would do so as well. Neurons have—are—their own logic. Neurons are compact and functional processing elements (although some neurobiologists would argue that single neurons are assemblies of several processing elements, or only components of a larger element).

In support of this view, recent findings suggest that neuromorphism may convey computational advantages on neural nets. One stimulus toward a synthesis of neural networks and neurobiology has been the observation that properties of hidden units of network models trained on neurobiological problems resemble biological interneurons engaged in processing similar information in the nervous system. Such correspondence is described in chapter 5 by Lockery and Sejnowski, and has been reported by Zipser and Andersen (1988a) and others. Zipser (1992) argues that this correspondence is in part to be expected from the class of *identification models* to which successful networks belong, but he also notes that such correspondence also depends on the realism of the model's components. Making neural network models more neuromorphic, therefore, would be expected to at least maintain their processing power intact, and at best increase their capability and at the same time make their algorithms and internal representations more relevant to neurobiology.

Neuromorphism may be more than biologically satisfying for third-generation networks; it may be an essential feature of their expanded capability. A recent analysis by Geman et al. (1992) suggests that

learning purely by example is not feasible for large classes of nonparametric problems, and that neural networks will be able to do no more than fine-tune or refine solutions that are partially hard-wired. In light of this analysis, it may be possible to view neuromorphism as imposing sufficient restraints or limits on a neural network—biological or artificial—to provide the structure biasing the network toward the solutions needed. In this view, neuromorphism would channel or target the solution space, itself helping to define the classes of problems that neural networks solve.

NEURAL NETWORKS: THE EXCITEMENT PERSISTS

Many of the concerns and goals of those who would create *neural* neural networks were laid out concisely by Francis Crick, who reviewed and summarized the mutual inconsistencies among neurobiology, neural networks, and psychology in a seminal *Nature* commentary (Crick 1989). The areas of concern he noted include:

- the neurobiological plausibility of backpropagation,
- the role of synapses using NMDA receptors,
- dynamic synaptic strengths,
- optimization of scaling and the number of processing units,
- dual excitatory and inhibitory synaptic outputs,
- the dual nature of network models as models of the brain and as generators of computationally useful algorithms, and
- the need to relate real neurobiological data to models and to model classes.

Others have contributed to this call for increased neuromorphism. Selverston (1988) criticized connectionism for its disregard of cellular, synaptic, and neuromodulatory properties, noting that truly neuromorphic models would encompass specific characteristics, not only of particular neuronal types, but also of individual and distinctive neurons. Getting (1989) reviewed several complex properties of biological neural networks and stressed the distinction between anatomical and functional connectivity of neuronal assemblies. A similar analysis was carried out by Tank (1989), who commented that the functions that a circuit implements determine the nature of the computation that it performs. Although the functions of a circuit are in part a consequence of the neurobiology of its components and assembly, several physical mechanisms or *constructions* of varying neuromorphism are capable of implementing the same function or *computation*. Clark et al. (1992) expanded the advantages of neuromorphism by suggesting additional, extrasynaptic, sites for neurobiological plasticity.

This volume is in part an examination of, a guide to, and a suggested solution for many of the areas noted by these authors.

THE NEUROBIOLOGY OF NEURAL NETWORKS AND THE THIRD GENERATION

The Neurobiology of Neural Networks is thus a manifesto. It calls for an intersection of contemporary neurobiology and artificial neural network models, to the mutual advantage of each. Exploiting this advantage requires not just a wary and distant appreciation by each field of the virtues and pitfalls of the other, but the realization that each area of inquiry is best served by a careful and selective convergence of information and point of view.

Each of the chapters in this volume attempts such a convergence. Furthermore, each approaches the neurobiology of neural networks in one of two ways. In the first part of the volume, chapters 2 through 5 take a bottom-up approach, examining features of synaptic transmission and suggesting ways in which each can be used to make neural network models simultaneously more neuromorphic and more computationally powerful. Chapter 5 forms a bridge to the second part of the volume, consisting of chapters 6 through 7, each of which presents neural networks modeling basic functions of the nervous system: motor control and sensorimotor integration. These functions were selected rather than the more common sensory processing or pattern recognition in part because they have received less attention, in part because they represent more fundamental, phylogenetically older systems. Reflecting the goal-directed nature of these neurobiological functions, each of these chapters partially or entirely takes a top-down approach. The emphasis in these latter chapters is on dynamic recurrent networks, with feedback and time-dependent enhancements partially consistent with neurobiology.

Early chapters relate biological synaptic strength to the synaptic weights characterizing neural network models. For example, in chapter 3, I draw explicit links between biology and models, both for synaptic weights themselves and plausible mechanisms for one popular learning rule. In chapter 4, Baxter and Byrne extend this analysis for classes of learning rules. The importance of synaptic weights has been stressed by Douglas and Martin (1991), who note:

For the present it seems that the greatest immediate gains in understanding the collective action of neurones will come from examination of the basic organization of neural connections and their approximate functional weights. It is exactly to this aspect that computational neuroscientists have paid the least attention.

Although my co-authors and I would agree as to the importance of synaptic weights in developing third-generation neural nets, this is not the only area of importance, and not the only one on which we concentrate.

Two Principles of Brain Organization: A Challenge for Artificial Neural Networks

In the next chapter, Stevens selects two common features of biological neural networks and shows how each requires modification from standard artificial neural network formalisms. These two are presented, as his title indicates, as a challenge to future generations of networks. Stevens presents data showing that as the cerebral cortex scales with evolution, its degree of interconnectivity implies specific patterns of modularity, rather than indiscriminate linkages. He cautions that network models will need to incorporate this modularity in order to encapsulate cortical algorithms. Again, this aspect of neuromorphism may enhance processing power as well as biological significance. Zipser (1989) found that one technique for forcing modularity, subgrouping, enhanced the performance of a recurrent network model by an order of magnitude. Similar increases in performance were obtained by introducing selective, modular rules for error propagation (Ballard 1990).

Stevens also examines the consequences for neural network models of the operation of synapses utilizing the recently characterized NMDA channel. Although it is recognized that these synapses introduce short-term memory, Stevens points out that in addition they require synaptic weights themselves to be dynamically reprogrammable, depending from moment to moment upon the pattern of synaptic input. Again, this feature will require modification of the present generation of network models. In a related context, Durbin and Rumelhart (1989) have developed a sum-of-products network unit that they suggest is partially analogous to NMDA synapses and therefore biologically plausible. Also supporting Stevens' challenge, Bernander et al. (1991) find that neuronal *integration* as well can be altered dynamically by strong synaptic input reflecting changed global network activity.

Static Determinants of Synaptic Strength

I begin chapter 3 with a short review for readers who are more familiar with one, rather than both, of the parent fields. I present neuronal function, network architecture, and synaptic strength and its modification from dual perspectives: that of neurobiology and that of the formalism of second-generation connectionist networks. Using a biological network of identified neurons, I focus on details of presynaptic and postsynaptic mechanisms determining static values of synaptic efficacy which result from individual experience. These data and related findings from hippocampal synapses and others are used to derive a hypothetical yet plausible biological mechanism for implementing some features of the back-propagation algorithm used in many second-generation artificial networks.

Learning Rules from Neurobiology

In chapter 4, Baxter and Byrne link neurobiology and network models by examining learning rules for modification of synaptic weights and relating each to observed synaptic plasticity. Recent investigations in hippocampal, cerebellar, and invertebrate networks are compared in a critical survey synthesizing interneuronal biophysics and postulated learning rules. These authors report evidence for many forms of associative and nonassociative learning rules and analyze each both neurobiologically and mathematically. Their most detailed case emphasizes contemporary discoveries of mechanisms for activity-dependent neuromodulation.

Realistic Network Models of Distributed Processing in the Leech

The work reported by Lockery and Sejnowski in chapter 5 provides two extensions toward the third generation of neural network models. In order to model an invertebrate reflex, they develop a network that incorporates components of conventional compartmental models into the framework of an artificial neural network employing back-propagation. This third-generation approach uses increased neuromorphism simultaneously to extend the restricted formalism of present-generation networks and to specify neurobiologically some otherwise-unrestricted synaptic weights. From the viewpoint of conventional compartmental analyses, this synthesis also allows neural network techniques to supplement these neurobiological models by providing higher-level representations that hide uncertain details of particular components. When the network is trained using biological data, hidden layer unit properties of this network predict the behavior of biological interneurons, suggesting that the algorithms developed by the network model (if not the optimization methods for deriving the algorithms) resemble those of the equivalent biological network.

Neural and Peripheral Dynamics as Determinants of Patterned Motor Behavior

In chapter 6, Chiel and Beer report development of a neural network controller based on neurobiological data and constraints, exploration of the properties of this controller, and its installation into a robot to examine real-world interactions. They further remind us that biological neural network controllers operate in a closed-loop configuration (as most network models do not, except in training or other monitored cases), and show the consequences of model interaction with peripheral effectors and the environment. Their controller shows complex behavior, with some patterns dependent on interactions with the periphery and others apparently independent. In this real-world case, second-generation error-minimization techniques are not used for setting

synaptic weights, which are instead varied systematically in order to reveal the strengths necessary to implement a range of behaviors.

Dynamic Neural Network Models of Sensorimotor Behavior

In the final chapter, Fetz describes classes of neurobiological and other problems to which dynamic recurrent networks have been applied, with emphasis on sensorimotor integration and on primate and other mammalian nervous systems. He reminds us that the inadequacy of single-unit recordings for deriving neuronal algorithms reinforces the importance of a neural network approach. In addition, he briefly analyzes several areas of intersection between neurobiology and network models, including time-dependent and short-term memory, processing of discrete action potentials, and neuronal units with nondifferentiable input/output functions.

BEYOND THE NEUROBIOLOGY OF NEURAL NETWORKS

As an introduction, this chapter needs no conclusion; as a guide and a call to future research, it should end only with the development of third-generation networks. Although this volume attempts to cover major areas in which third-generation neural networks could be made more neuromorphic, neurobiology is complex, and there is much remaining. Several additional extensions are likely to enhance future networks.

Although several groups including Ackley et al. (1985) have used stochastic methods to assist convergence, neural network models have only recently begun to incorporate the stochastic nature of neurotransmitter release and its effects on neuronal firing patterns and network computation (Bulsara et al. 1992; Burnod and Korn 1989; Burton and Mpitsos 1992; Smith 1992).

Another biologically suspect characteristic of artificial network models is the separation between the training or acquisition phase and the production or test phase in which the converged and appropriate weights are used for information processing. Even prior neurobiological distinctions between development and adult learning are also becoming blurred. As an example, Kandel and O'Dell (1992) review recent work on plasticity suggesting "a temporal continuity between development and learning." They note further, drawing on the work of several laboratories, that the "best candidates for mediating [developmental] processes" are those such as NMDA receptors that mediate some forms of adult learning.

The second-generation segregation of elementary model units into layers represents partial recognition of the importance of connectivity, but is more rigidly arbitrary than neurobiology. Anatomical and physiological circuitry is known for several neural assemblies, and is beginning to be used in place of massive interconnectivity as the third generation of artificial neural network models is developed.

Finally, an engineering appreciation of neuromorphism has proven useful as a guide to the design of silicon-based electronic components (Faggin and Mead 1990; Mahowald and Douglas 1991). These neuronal analogs are capable of operation at speeds 5 to 6 orders of magnitude faster than their biological counterparts; it is not yet clear if this is an advantage or a deficit.

ACKNOWLEDGMENT

I thank the Office of Naval Research and the National Institutes of Health for their support through grants N00014-90-J-1490 and NS-11555.

2

Two Principles of Brain Organization: A Challenge for Artificial Neural Networks

Charles F. Stevens

One use for artificial neural networks is to give insights into brain function by providing tests of possible computational principles the actual neural circuits might employ. Of course, the artificial network need not correspond closely to an actual network in order for this approach to be useful. That is, the artificial network might use the same algorithm as does the brain, but it might implement this algorithm in a quite different way. Many would agree, however, that artificial neural networks are most valuable to neurobiology when they are more than metaphorical, when their circuits capture the functional essence of the actual brain network. The goal of this chapter is to describe two classes of constraints on artificial neural networks provided by actual networks. Other chapters in this volume discuss additional principles and constraints of biological nervous systems. Each of these constraints poses a challenge for the computational neuroscientist who wishes to develop accurate models of neuronal circuits.

The first of the two classes of constraints has to do with scaling properties of actual neuronal circuits in the mammalian cortex. The second incorporates properties of a recently characterized dual-mode synaptic function that differs fundamentally from the type of synaptic connections generally used in artificial neural networks.

SCALING PROPERTIES OF CORTEX

The use of animals in experimental neuroscience typically involves the assumption, implicit or explicit, that conclusions derived from the study of one species will, at least for the most part, apply to others. The obvious basis for this assumption is the mechanism of evolution in which structures are modified gradually, if at all, from one species to another. Abundant experimental evidence justifies neuroscientists' belief that all mammalian brains share the same basic design. General principles and even many details discovered from studies on, for example, cats, are found to apply to rodents or primates.

Mammalian brains vary in size by about four orders of magnitude, from less than a gram for some shrews to nearly 10 kg for large whales

(for examples, see references cited in Prothero and Sundsten 1984). The primary visual cortex (V1) of primates ranges in size over about two orders of magnitude (Frahm et al. 1984), yet the difference between small and large visual areas in their function and pattern of neuronal interconnections are very minor. What scaling principle does the brain use for neuronal computers of such varying size? More specifically, how does the number of neuronal interconnections and the distance over which the interconnecting axons spread vary as the size of the neuronal computer is increased (Ringo 1991; Stevens 1989)? One might imagine, for example, that larger numbers of computing elements, or neurons, require more interconnections between the individual units. The great majority of connections between neurons are intrinsic, so perhaps a constant fraction of all neurons in a particular cortical region, V1, say, communicate. Were this true, then the average number of synaptic contacts per unit would have to increase in proportion to cortical size. Alternatively, it could be that cortex is a modular structure, so larger cortices have increased numbers of modules but the interconnectedness of individual units within a module is independent of cortical size. Finally, the brain might use a design principle in which the number of computational elements is traded for the complexity of each element, so that the number of connections each neuron received would decrease as cortical size increases. Whatever the answer, an artificial neural network that is claimed to implement literally a biological neural network would have to use the same scaling principle as the brain.

Anatomical data to determine how interconnections scale are in the literature, but extracting the information from the available findings requires an appropriate formulation of the problem (Stevens 1989). Experimental observations provided by the references listed in Stevens (1989) may be used to derive three principles: (1) Within the cortex, the density ρ of synapses is nearly constant across cortical regions and layers, and has an approximate value $\rho \approx 10^9$ synapses/mm³. (2) The number of neurons m underlying each unit area of cortical surface is approximately constant for all cortical regions and is about 150,000 neurons/mm². The exception to this rule is primate V1, where the number is about twice the value for other cortical regions. (3) The thickness T of cortex increases very slowly with the cortical surface area A such that $T \propto A^b$, where $b < 0.1$.

These experimentally observed principles may be used to derive the following theory of neuronal interconnectedness. The size of some identifiable cortical region will be specified by N , the number of neurons in that region. For example, our visual cortex contains about 3 billion neurons, whereas that of the squirrel monkey contains about half a billion. Let N_r be the number of neurons in the cortex of some species taken as a reference. If we choose the squirrel monkey as a reference, $N_r = 0.5 \times 10^9$. The average convergence, or number of synapses each neuron receives, s , will in general depend on the cortical size N , so that

$s = s(N)$. Because all primate V1 cortices are believed to follow the same design, the average number of synapses per neuron in a given cortex of size N is related to the number in the reference brain:

$$s(aN_r) = f(a) s(N_r) , \quad (2.1)$$

where $a = N/N_r$ is the scale factor relating the sizes of the two homologous cortical regions and $f(a)$ is some function of the scale factor a . The key idea here is that $f(a)$ is independent of absolute brain size; this is equivalent to saying that, because of the way evolution works, any brain serves equally well as the reference. The functional equation above is well known to require that $f(a)$ be a power function:

$$f(a) \sim a^b \quad (2.2)$$

for some constant b . This means that:

$$s(N) = s(N_r) (N/N_r)^b . \quad (2.3)$$

The average number of synapses per neuron thus varies as a power law with brain scale a , so that the exponent in the power law gives a measure of interconnectedness: if $b = 0$, then the number of synapses per neuron is independent of cortical size, whereas if $b = 1$, then each neuron is connected to a constant fraction of the other neurons in the cortex.

Some simple substitutions permit b to be estimated from experimental data. The average number of synapses per neuron $s(N)$ for a cortical region with surface area A and thickness T , is by definition (*total number of synapses/number of neurons*):

$$s(N) = \rho AT / mA = T\rho / m ; \quad (2.4)$$

ρ is the synaptic density ($10^9/\text{mm}^3$) and $m \approx 150,000$ is the number of neurons underlying a square mm of cortical surface ($m \approx 300,000$ for V1). The number of neurons in a cortical region is just:

$$N = mA. \quad (2.5)$$

Substituting these relationships into the power law derived earlier, we find:

$$\frac{T\rho}{m} = \frac{T_r\rho}{m} \left(\frac{mA}{mA_r} \right)^b ; \quad (2.6)$$

here T_r is the thickness, and A_r the area, of V1 in the reference brain. Canceling common factors, we find:

$$T = T_r \left(\frac{A}{A_r} \right)^b . \quad (2.7)$$

We know (see above) that this power law relating cortical thickness and surface area corresponds to what is found experimentally, and that b has a value less than about 0.1. Thus, the number of synapses per neuron increases by (at most, as we shall see) a factor of $10^{0.1} = 1.26$ for a tenfold increase in cortical size.

This value is, however, an overestimate. Unlike tissues other than brain—in which the cell size is independent of organ size—larger brains have larger neurons (Bekkers and Stevens 1990c; Bok 1959; Hardesty 1902). In general, the size of a neuronal cell body corresponds to the length of the axonal processes it must maintain: Betz cells and spinal motor neurons, the cells with the longest axons, are the largest whereas local interneurons are small. Larger brains, then, have neurons with longer axons and larger cell bodies. In addition, some, but not all, neurons in larger brains have larger diameter and longer dendrites (Bekkers and Stevens 1990c). All of these factors combine to make larger brains have thicker cortices. Because cortical surface area is approximately proportional to brain volume and homologous areas tend to occupy a constant fraction of the total cortex, cortices containing more neurons are thicker. The derivation of the earlier equation assumed that the size of the constituent elements are independent of brain size, but this is not strictly true. Thus the equation derived above actually places an upper limit on interconnectedness.

I conclude that neuronal interconnectedness is essentially independent of cortical size, an observation that suggests, but does not require, a modular construction of cortex. Relevant to this issue of modularity is the question of how far cortical connections reach: Does, for example, a neuron in a large cortex send information over the same absolute distance or over the same fraction of the cortex? For V1 we know that the entire field of vision is mapped in an orderly way over the entire cortical region. The preceding question can be reformulated: For a large visual cortex compared to a small one, are the connections maintained between corresponding parts of the visual field, or is the absolute length of axons that make up the neural circuits approximately preserved?

Most synaptic connections in a cortex are, as noted earlier, intrinsic and a significant fraction of the cortical volume in all layers (Gilbert and Wiesel 1983; LeVay 1988; Livingstone and Hubel 1984; Rockland 1983; Rockland and Lund 1982) is contributed by the axons that carry intracortical information. If the intracortical axon arbor span were maintained as a constant fraction of the cortical length, then the volume contributed by intrinsic axons would increase by a factor greater than $\sqrt{N}$. Because the cortical thickness hardly changes with surface area, a neuron's arbor must extend over a distance that is largely independent of cortical size, for otherwise the increased axonal volume would make the cortex much thicker in larger brains. This argument can be sharpened: Specific connections are maintained by axons that travel in defined cortical laminae, and the laminal thickness is mostly independent of cortical size (Rockel et al. 1980). Thus, all of the components of a circuit that travel in specific laminae have about the same absolute span in cortices of all sizes as would be expected for a modular design.

The arguments in the preceding paragraphs have been framed in terms of a particular set of homologous cortices, primate V1. Although

the various species devote different fractions of their total cortex to specific functions—some animals, like primates, have large visual systems, and other animals have relatively larger olfactory systems—on average most mammalian brains seem to follow quite comparable designs. The preceding arguments can be carried through without any change for total mammalian cortex, and the conclusions are unchanged for this larger universe (Stevens 1989).

In summary, mammalian cortex appears to have a modular design in which the degree and span of neuronal interconnectedness hardly varies when the size of neural computer ranges over four orders of magnitude. This fact provides a powerful constraint for artificial neural nets that are intended to mimic the operations of actual neuronal circuits.

CLASSES OF NEURAL NETWORKS

Artificial neural nets typically start with some rule governing connections (everything in one layer hooks to everything in the next, for example) and then the synaptic weights are adjusted with a suitable learning algorithm, such as backpropagation (Rumelhart et al. 1986a). The end result is a network whose computational properties depend on the input-output characteristics assigned to the individual units and the pattern of synaptic strengths that have resulted from training. An important assumption for this class of networks is that synaptic strength is fixed except as modified by the learning rule. That is, the synaptic weights at a particular time depend only on the training history and not, moment to moment, on the pattern of inputs to an individual unit. Neural computers of this sort will be designated *class one*. For these synapses, synaptic strength is relatively static; neurobiological mechanisms and correlates for class one synapses are discussed in chapters 3 and 4.

One can, however, imagine a neural computer for which the synaptic weights onto a particular unit are functions of *both* history and of current synaptic input activity that unit is receiving. In general, synaptic strength might be some function of the particular spatial pattern of inputs onto a unit that are currently active, but in a simple case strength of certain inputs would depend only, for example, in the total input activity. Neural computers of this sort will be referred to as *class two*. For these class two synapses, synaptic strength is thus dynamic rather than static. (Several different dynamic properties of neural networks are presented in chapters 5 and 7.)

Although the amount of current injected into a neuron by a synapse depends on the neuron's membrane potential, and thus on the activity of its other inputs, weights defined in terms of synaptic conductance are independent of other inputs the cell is receiving. Furthermore, even if weights are expressed in terms of synaptic currents (or of the resultant voltage changes), these weights become independent of voltage at the

limit of small deviations from the neuron's resting potential and they show a saturation phenomenon only for large depolarizations. Thus, for classical synapses, class one computers are a good approximation however synaptic weights are assigned (conductance, injected current, or membrane potential change). Networks that employ classical synapses are therefore in class one.

Over the past several years, neurobiologists have increasingly appreciated that many central excitatory synapses are dual functional (see, for example, Bekkers and Stevens 1989). That is, the postsynaptic response at these dual functional synapses results from the simultaneous activation of two broad classes of intrinsic membrane proteins generally called NMDA and non-NMDA receptors. The non-NMDA receptors subserve "normal" synaptic transmission: The response is a rapid transient and the conductance changes are independent of membrane potential. The NMDA receptors, however, have properties that place neural computers that use this receptor type in class two. We turn now to a discussion of NMDA receptor-mediated component of synaptic transmission.

NMDA receptor function differs in two significant regards from that of the non-NMDA receptors. First, the effects of a single activation persist for several hundreds of milliseconds, whereas the non-NMDA component is only a few milliseconds in duration. Thus, the NMDA receptors provide a form of transient memory for recent synaptic activation. Second, and more important, the conductance of the NMDA receptor channels depends jointly on the binding of glutamate and on the neuron's membrane potential (Mayer et al. 1984; Nowak et al. 1984). For this reason, the weight of this synaptic component depends on the activity of other inputs because these inputs determine the neuron's membrane potential.

The transient memory provided by the NMDA receptors is not the result of continuing presence of glutamate in the synaptic cleft. Rather, the glutamate that a receptor binds is retained at the receptor's binding site for the several hundreds of milliseconds of receptor activation. Neurons in the hippocampus, for example, tend to discharge impulses in epochs of activity that occur once about every one to two hundred milliseconds (the theta rhythm; see Bland 1986); NMDA receptors provide a memory that lasts about a theta rhythm cycle.

The mechanism through which NMDA receptor channel conductance depends on membrane potential is well understood (Jahr and Stevens 1990a, b). Glutamate binding opens the channel after a delay that averages five to ten milliseconds. As soon as the channels open, magnesium ions present in the extracellular fluid rapidly enter the open pore and block the current flow through the pore. The Mg^{2+} ions must move through a fraction of the membrane's electric field to reach their blocking location, and this means that the energy of the ion in its blocking site, and thus the blocking ion's dwell time at the site, depends on the membrane voltage. A well-tested quantitative theory for this process

relates the conductance of the channel population $g(V,t)$ to the membrane potential V and the time t by the equation (Jahr and Stevens 1990b):

$$g(V,t) = \frac{G(t)}{1 + 0.28 [\text{Mg}] e^{0.062V}} \quad (2.8)$$

where $[\text{Mg}]$ is the extracellular magnesium concentration (in mM/liter); $G(t)$ is the NMDA receptor channel conductance as a function of time t for the limiting case of large depolarizations where there is no magnesium block. This equation is an approximate one that is adequate for physiological Mg^{2+} concentrations. Note that the equation implies that the NMDA receptor component of synaptic potentials can be turned on and off by voltage changes resulting from other of the neuron's excitatory and inhibitory inputs for the time it takes $G(t)$ to decay back to zero. This time is, as noted above, one or two hundred milliseconds. Thus the transient memory of synaptic activation that is maintained by the NMDA receptors can be "read out" over a comparatively long time by other synaptic inputs.

The voltage-dependent Mg^{2+} block is a key characteristic of NMDA receptors, and it places the regions of the brain that use these receptors in the second class of neural networks. Not all neural computations rely on the NMDA receptor component of synaptic transmission, but some of them do (see Bekkers and Stevens 1990a). These computations, then, are carried out by the neuronal network operating in the class two mode, and the modeling of these computations by a class one network could not be literally correct.

One consequence of the activation of NMDA receptors in some areas of cortex is the subsequent alteration of static synaptic strength. In contrast to the computations performed by these class two elements, the mechanisms they utilize to change synaptic efficacy have parallels in artificial neural network models, and these will be evaluated in chapter 3.

CONCLUSIONS

In summary, two general features of the brain—its scaling properties and the fact that some of its computations are performed by a neuronal network operating in a class two mode—provide challenges for artificial neural networks that are supposed to operate in the same way as actual networks. Any neural network that is claimed to operate as does the cortex would have to scale correctly and would have to fall into the same computational class as the actual neuronal network. Of course, many insights about brain function can be gained by artificial networks that do not meet these criteria, and such networks could be useful for solving practical problems or for revealing general principles about other artificial networks. Still, the full computational power of the brain

cannot be modeled until artificial networks are developed that scale appropriately and that carry out computations in both of the modes that characterize brain synapses.

ACKNOWLEDGMENT

I gratefully acknowledge support from NIH-NINDS Research Grant NS12961 and from the Howard Hughes Medical Institute.

Daniel Gardner

The ability of biological nervous systems to process, classify, and act on information derives not only from the interneuronal connectivity, but also from the often complex properties of each synapse in the network. These synaptic properties need not be fixed, but are often plastic, changing in response to rules for modifying synaptic efficacy that may incorporate signals both intrinsic and extrinsic. Even properties of apparently static synapses are likely to reflect the past influence of plastic modulation.

This chapter concentrates on neurobiological determinants of synaptic function that are likely to be relevant to interneuronal information processing. The data presented can be used to evaluate rules or algorithms used to alter such properties as synaptic strength, and to relate these to function. By examining differences in synaptic efficacy, one can determine in part how these differences are effected, and if they reflect presynaptic or postsynaptic control. It is also possible to ask if synaptic strengths are independently alterable at different sites of the same neuron: at separate presynaptic terminals or separate postsynaptic specializations. Although related questions have been investigated in such mammalian preparations as hippocampus and spinal cord, the major findings presented in this chapter utilize an invertebrate preparation: the buccal ganglia of the marine mollusc *Aplysia californica*. This preparation has two advantages for these studies. First is an experimental accessibility which permits reliable low-noise voltage clamping of several neurons simultaneously. The second derives from the network properties of a defined set of interconnections.

Network properties are of particular importance because nervous systems are composed of interconnected arrays of neurons with similar, but not identical, properties. Many analyses describe a nervous system principally in terms of the interconnectivity of the neuronal assembly. Others offer an additional viewpoint, describing properties of classes of neurons and distinguishing between different classes. However, even within a class of similar neurons, individual cells can differ from one another. Only some preparations permit an additional level of analysis, examining specific features that distinguish neurons as individuals, rather than combining general properties that are common to classes of cells.

As a result, it becomes possible to recognize individual neurons and to *identify* them, using criteria such as shape, size, position, electrophysiology, and connectivity. The use of identified cells extends analysis to include the incremental assembly of libraries of specific information characterizing each cell, by repeated study of the same cell in a series of preparations. Network as well as cellular analyses may be extended similarly, by elucidation of specific network properties that characterize interconnections of individual cells, rather than cell types. One such network property is the efficacy of the synaptic connections made between individual neurons.

The ability to study individual cells and the efficacy of individual synapses in relation to their place in a neuronal network characterizes only a few biological nervous systems, but has widespread significance not only for biological nervous systems in general, but also for neural network models. These arrays of artificial cell-like elements use various computational algorithms, or learning rules, to adjust the strength of information transfer between cells. The computational ability of these connectionist networks is an emergent property of variable synaptic weights and of local processes that change synaptic weight (for reviews, see Anderson and Rosenfeld 1988; Anderson, Pellionisz, and Rosenfeld 1990; McClelland and Rumelhart 1986; Rumelhart and McClelland 1986).

In common with other papers in this volume, this chapter compares biological nervous systems to artificial neural network models. In particular, I attempt to relate biophysical properties of synapses to requirements for adaptive alteration of synaptic strength in the tuning of artificial network models. The parallels extend both ways: Mechanisms for synaptic variability not only have relevance for neuroscience, but also may aid the refinement of neurocomputational models. Both neurobiology and neural networks can contribute to, and each may benefit from, the exploration of mechanisms available to determine synaptic strengths within a network processing neural information.

Part I of this chapter begins by reviewing basic features of connectionist network models from a neurobiological perspective. I then present the preparation used in this study, the buccal ganglia of *Aplysia*, both as a biological nervous system and also as an analog of neural network models, discussing similarities and differences between model and ganglion. The use of identified cells with a stereotyped wiring diagram permits analysis of these neurons as if they were elements of a connectionist network. Here, the major parallel drawn between biological and artificial networks relates to differing and variable synaptic strengths.

Part II explores the neurobiology of common learning rules for error minimization in multilayer connectionist neural networks, rules which require reverse information transmission across model synapses. Utilizing inherent network properties of the *Aplysia* buccal ganglia, I show that in one particular network in this preparation, synaptic strength is specified by postsynaptic neurons, yet controlled by varying presynap-

tic transmitter release. From these data, the existence of retrograde information flow can be postulated in this biological preparation. The possible role of each of three classes of mechanisms that might yield the results presented is examined. The chapter then reviews the evidence for reverse synaptic information flow in other preparations. This includes dynamic retrograde modulation at hippocampal synapses, and static postsynaptic specification of presynaptic release elsewhere. To more closely relate these neurobiological data to neural network models, I discuss classes of learning rules that might yield the results presented. The chapter closes by showing that similar neurobiological mechanisms could be invoked to implement the requirement for reverse information transmission required by connectionist models, suggesting that such models need not be dismissed as neurobiologically implausible.

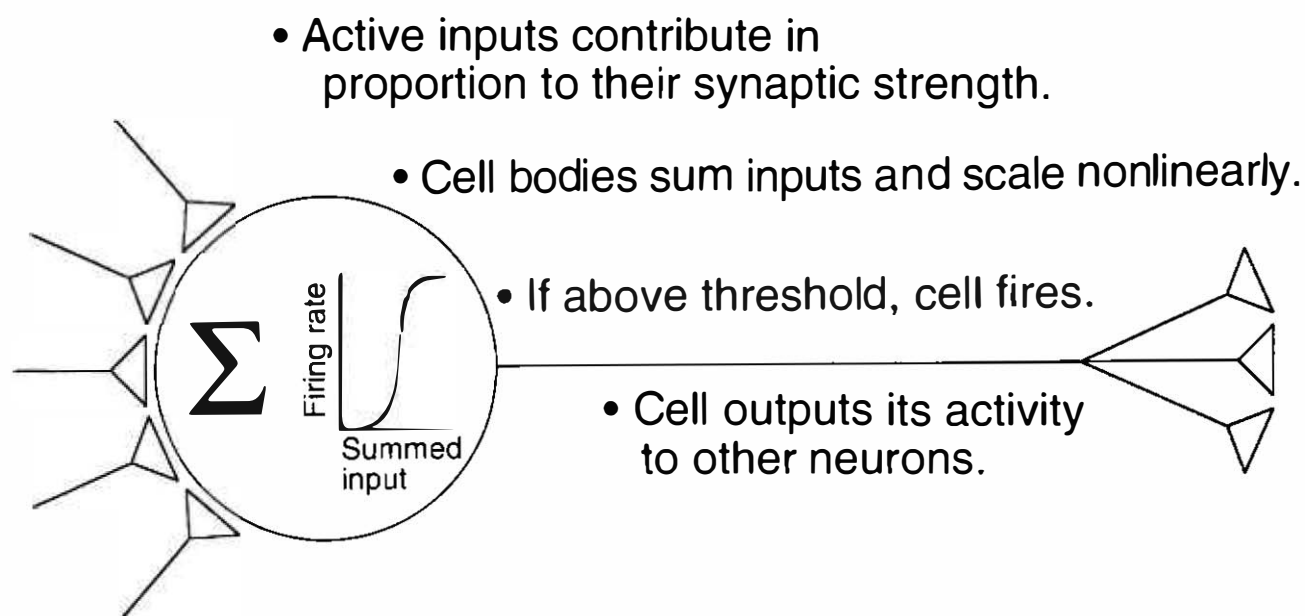
I. A Biological Neural Network

THE NEUROPHYSIOLOGY OF NEURAL NETWORK MODELS

Neural network models, although not directly derived from the nervous system of any real organism, are composed of elements, often called neurons, interconnected by synapse-like mechanisms, and are often applied to problems in information processing similar to those carried out by parts of the brain. To a neurobiologist, it is legitimate to study artificial network models as if they were biological nervous systems and then to assess the relevance of these models to assemblies of real cells. Such analyses can reveal if the two have features in common, if their similar terminology is derived more causally than casually, and if one can predict features of the other.

The particular model networks I will draw analogies to are classic connectionist networks, implementable with electronic components but universally simulated on a computer. These models, with only minor differences among them, can be characterized by three components: the elements of the model (neurons and synapses); their architecture (here, equated with the wiring diagram), which incorporates extensive, but not indiscriminate, convergence and divergence, most often in three or more layers; and finally the learning rules for modifying synaptic strength. Of course, both elements and architecture represent simplifications, disregarding many significant aspects of neuronal cell biology and interconnectivity.

Neurobiological examination of the components of neural network models begins with the individual processing elements. Figure 3.1 presents one of these simple units and some of its properties. This model element, or *neuron*, receives input from several other neurons, as shown by triangles synapsing on its cell body. Each input is characterized by a synaptic strength, or *weight*. Inputs from active neurons contribute

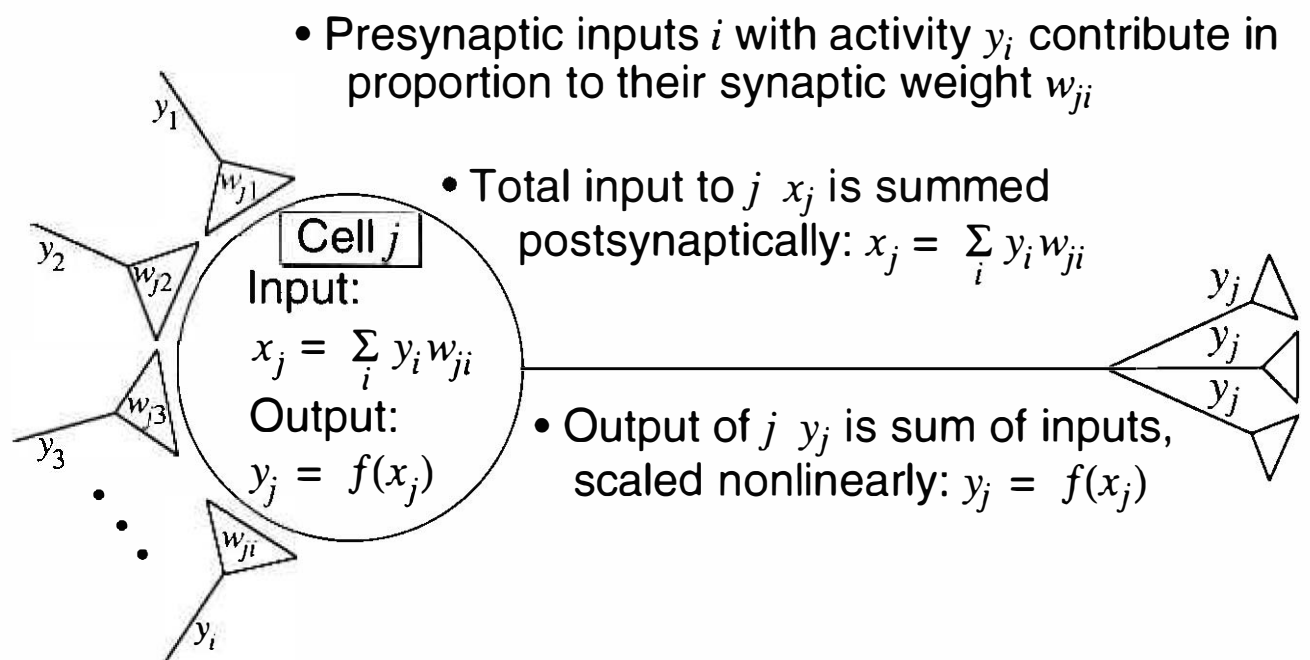


Neurobiology of a Model Neuron

Figure 3.1 Neurobiological description of a hypothetical neuron incorporating simplifications common to elements of neural network models. Anatomically, the cell body and all postsynaptic specializations are indicated by a circle, with presynaptic specializations encapsulated in multiple triangular terminations to an axon. Functionally, inputs contribute according to their activity, proportional to their synaptic strength. Cell bodies sum the input and scale it, as shown schematically, to produce an index of neural activity. This activity is in turn transmitted differentially to different higher-order neurons, again proportional to the synaptic strength of the individual connections. Compare to figure 3.2.

according to activity of the input unit, scaled or modulated by the strength of the synapse connecting the input. The activity of each unit can vary with time, and inputs from neurons that are inactive momentarily do not contribute to the model element on which they synapse. The cell body has two functions. First, it sums all the active inputs, each proportional to the strength of the corresponding synapse. Second, the cell body scales this input, most often by a nonlinear transformation, yielding an activity for the model neuron. This level of activity is then passed on to other cells at synaptic outputs, shown on the right in figure 3.1. Inhibitory inputs can be represented by negative weights. The nonlinear scaling of summed input, which is shown as a sigmoid function in the figure, often incorporates a threshold. The scaling function is discussed further in chapter 7.

The same model element can be viewed more formally, in terms of the defining equations (figure 3.2). In this common formalism, derived from Rumelhart et al. (1986a), cells in layers are characterized by indices $i, j, k, \dots$, synaptic weight by w , input activity by x , and scaled output activity by y . Because activity levels are uniform throughout a cell, x and y carry only a single subscript, indexed by cell, but synaptic strengths are indexed by both postsynaptic and presynaptic cell (in that order), so that the connection from cell i to cell j has weight w_{ji} . Each of these values is related, as shown in figure 3.2. The input activity to a cell j is the sum of scaled outputs of presynaptic cells i :



Network Equations of a Model Neuron

Figure 3.2 Neural network description of a neuron-like model element in the j th layer of a multilayer model. The sum of inputs is shown as x , scaled output as y , and synaptic weight from the i th to the j th cell as w_{ji} , following the notation of Rumelhart et al. (1986a). The synaptic input to cell j from cell i is $y_i w_{ji}$. The total synaptic input to j , x_j , is the linear summation of $y_i w_{ji}$. The output of cell j , y_j , is a nonlinear scaling by f of the total synaptic input. Compare to figure 3.1.

$$x_j = \sum_i y_i w_{ji} ; \quad (3.1)$$

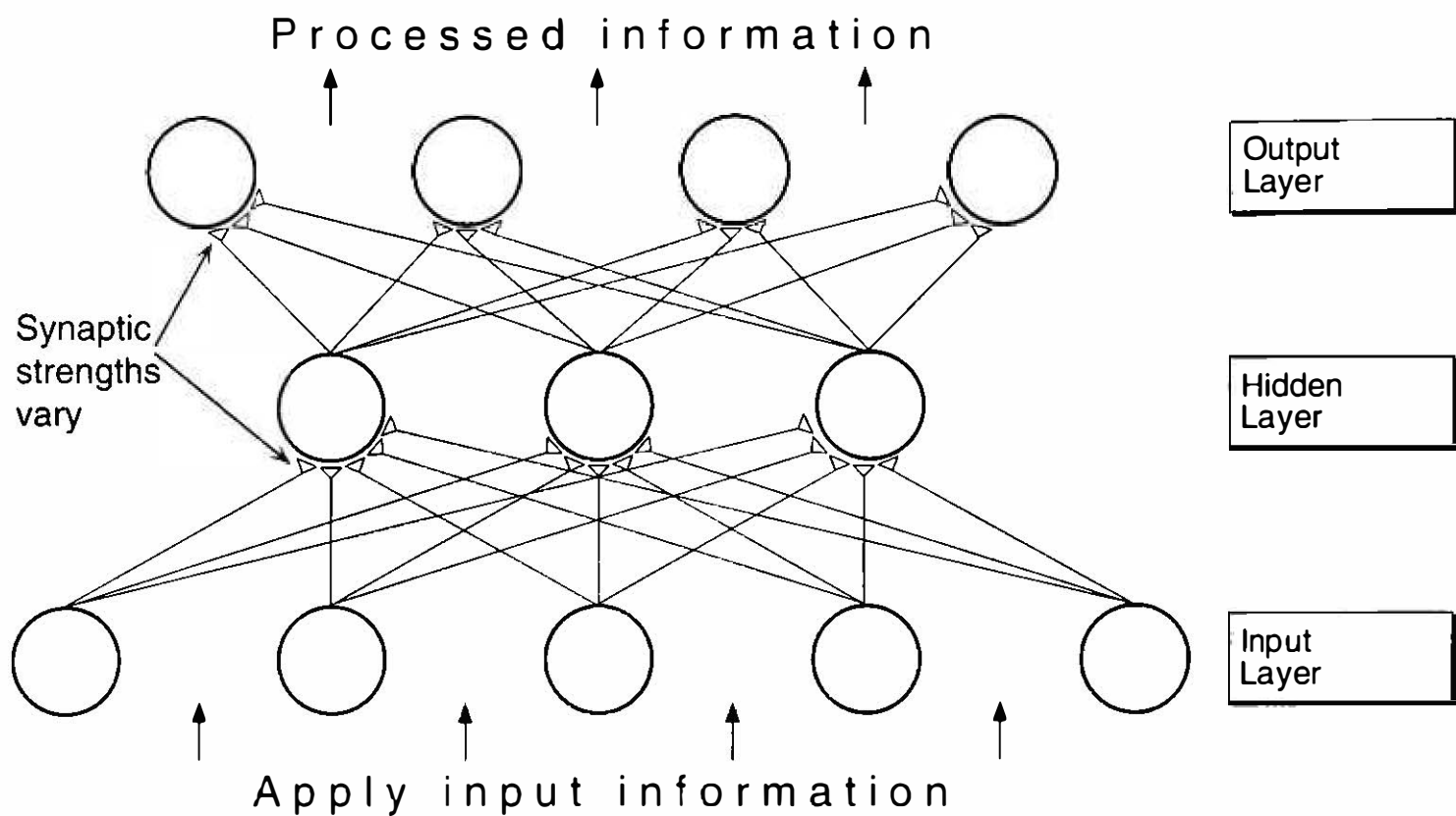
the output of cell j is a function of its input:

$$y_j = f(x_j) . \quad (3.2)$$

In order to strengthen the analogy to be drawn below, the synaptic weights are drawn as if they were local to individual presynaptic processes. Although not explicitly shown, each of these values is also a function of time; the function and design of most models is such that values of w vary more slowly than x or y .

The second component of a neural network model is its architecture, including the wiring diagram for its interconnection. Although it would be possible to massively interconnect all elements of a network in an indiscriminate nerve net, this would neither be consistent with most biological nervous systems, nor be a computationally effective way to process information. Instead, model neurons are arranged in layers.

Useful networks have three or more layers (figure 3.3). In this feedforward scheme, the input layer receives from external sources the information to be processed by the network. All the neurons in this layer send output to each of the cells in the next, or hidden, layer. Each of these hidden layer cells, in turn, sends its output to either other hidden layers, not shown in this example, or to the output layer, where the network output represents an appropriately transformed correspondence to the input. In a typical example, the input layer may be thought of as a sensorium,



Model Neurons are Arrayed in Layers

Figure 3.3 Schematic diagram of multilayer organization of typical neural network models. For clarity, only three layers are shown, the minimum number to solve several classes of interesting problems. The network is composed of the elemental processing units described in figure 3.2: neurons that sum and scale inputs according to a uniform formalism, and synapses that vary in weight or strength. The processing power of the network derives from combinations of these synaptic strengths. Networks are characterized by massive feedforward connectivity from each cell in a given layer to every neuron in the next higher-order layer. Information thus proceeds from the input or sensory layer through one or more hidden layers to the output layer.

perhaps a neural retina, and the output layer as composed of cells whose differing activities can be used to distinguish between specific inputs, such as faces or letters of the alphabet. The middle layers are conventionally described as hidden because unlike the input or output layers, the code for activity of these neurons is not defined. Just as in the case of single-unit recording in nervous systems, it is not possible to correlate activity of these units with real-world events without careful analysis. Often in networks transforming sensory information, these hidden units develop into feature detectors. (Extensions to this basic formalism, including such additional features as dynamic elements and recurrent or feedback connections, are discussed in chapters 5 and 7.)

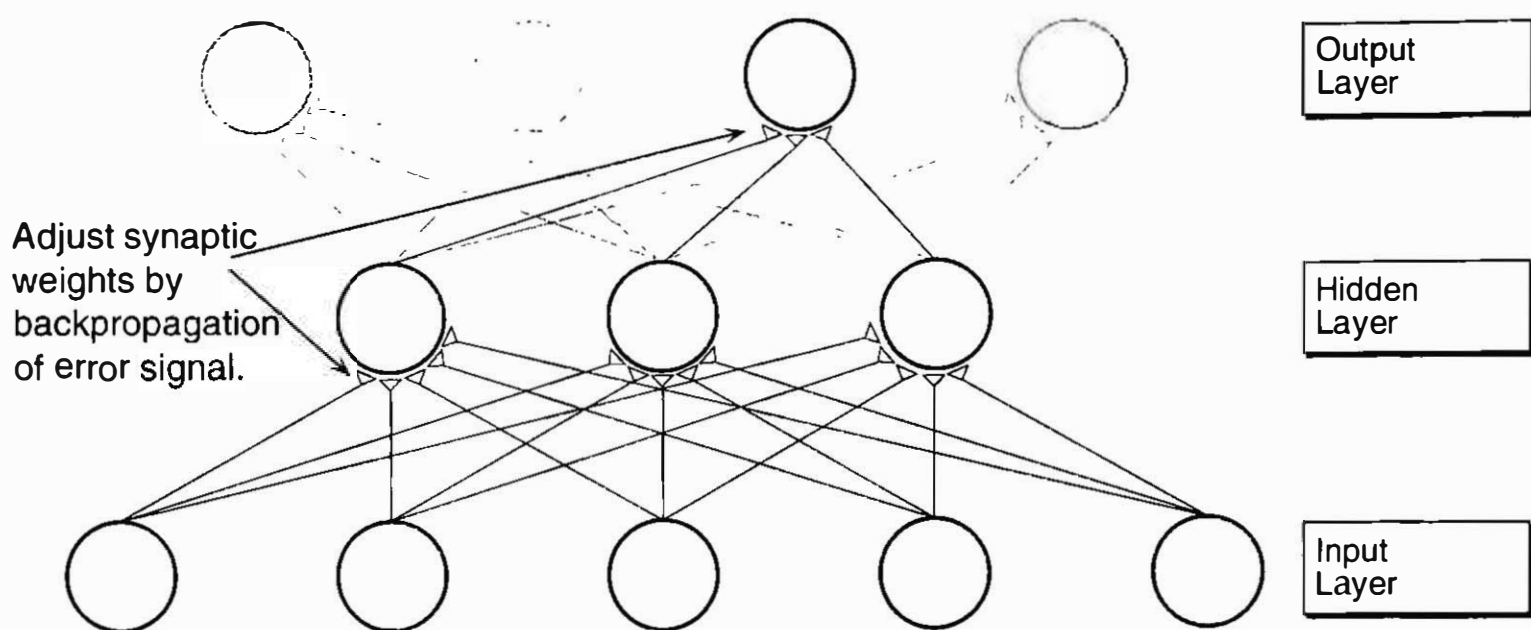
While typical network models may differ in the number of elements, in the class of inputs applied, and in the significance of the output, many practical networks conform closely to this scheme. Model construction often involves ad-hoc optimization of the scaling function $f(x)$, the number of cells and even the number of layers; once the network is designed, however, its anatomy is fixed. The number of cells and the wiring diagram do not change with learning or training. Similarly, properties of cellular elements are fixed and usually uniform through-

out the network. What defines a network, discriminates one similar network from another, and distinguishes a network that is successful at its information processing task from one that is not, is the set of strengths of the synapses that interconnect the cells: the synaptic weights between elements.

These synaptic weights, and the rules for varying them, form the third component of neural network models. Network models are typically constructed to solve a task in information processing that requires a distributed algorithm, or for which the details of the appropriate transform are unclear. Consequently, when a multilayer model is first constructed, appropriate values of synaptic weights are unknown, and they are set arbitrarily and randomly. If the network is then provided with an input, the input is processed through layers of neurons interconnected by inappropriately valued synapses, producing a random and usually improper output. The untrained or naive network is thus unable to perform the information processing task. Because the anatomy and cellular properties are fixed, the only way the network can be made more successful is to change its synaptic weights in very particular ways, raising some and lowering others. Manual adjustment of weights by the designer could improve network performance, but the number of weights in a typical network, as well as the obscurity of distributed algorithms, makes this method impractical.

Instead, the use of learning rules permits networks to modify their own synaptic weights so as to partially optimize the solution. In effect, the network adjusts its own synaptic strengths; it learns by doing. Useful networks provide appropriate transformations not just for one input pattern, but for large sets of partially related inputs. In a typical training procedure, a network is presented with paired patterns, consisting of individual input patterns and the desired output, by an external teacher. Sets of many inputs, covering the range of patterns for which processing is expected, each paired with the appropriate output, are presented repetitively. Training utilizes learning rules to modify synaptic weights so as to optimize network performance over the set of inputs used for training. After training, and depending on the task, the network model can successfully process similar inputs that were not part of the training set, analogous to generalization.

The most popular, and one of the most effective, learning rules is backpropagation; extensive descriptions of its techniques, antecedents, significance and utility are provided elsewhere (for examples, see McClelland and Rumelhart 1986; Rumelhart and McClelland 1986; Rumelhart et al. 1986a; Zipser and Rumelhart 1990). In common with several other learning rules, it relies on minimization of error. For each output unit of a network, the error is related to the difference between the observed and the desired activity of the unit, and varies with each input pattern. If the network were performing as we wished, the input



Error minimization requires propagation of components of the error signal from postsynaptic to presynaptic elements of each synapse in order to modify synaptic weights.

Synaptic Strength is Adjusted by Backpropagation

Figure 3.4 Neural network models adjust synaptic weight by backpropagation. During training, network models compare outputs produced by particular inputs to the desired outputs provided by an external teacher. The difference, or error signal, is propagated backward from postsynaptic to presynaptic elements in order to adjust synaptic weight so as to minimize the error, and also to provide a proportional component of the error signal to lower-order cells. Error signals from the output layer cell indicated will contribute to error signals in each of the other cells shown in black, and to adjusting synaptic weight of each of the synapses similarly shown in black. Analyses of cellular mechanisms contributing to backpropagation are provided in figures 3.7 and 3.8.

signal would be processed by the weights of the synapses it passes through to become the output, and there would be no error.

Optimizing network performance consists therefore in minimizing the error, not just for a single output, but for all output units, and not just for a single input/output pattern, but for the entire set. Analysis of the error values is used to adjust synaptic weights so as to make the error as small as possible (figure 3.4). In backpropagation, the error signal is propagated backward through the synapses that connect outputs with the inputs, and then each synaptic weight is changed—made stronger or weaker—so as to make the error as small as possible. The process is recursive in several ways. The scheme for error propagation itself relies on recursion to compute the error signal for layers feeding into the output layer. Also, sets of input patterns are presented repetitively until synaptic weights converge and the total error reaches a minimum. (This learning rule scheme requires an external teacher to provide the desired output and thus permit the error to be computed. There exist as well schemes that do not require a teacher, with comparable but somewhat different properties; some are discussed briefly in part II and in other

chapters as well. Chapter 4 includes an extensive discussion of neurobiological analogs to learning rules.)

The neurobiological relevance of two features of these connectionist networks is discussed in this chapter. First, there is often more than one way for a network to design itself to solve a particular problem. If one starts with not one but several structurally identical neural networks with initially randomized weights, and trains each of them on a slightly different set of examples of input and output patterns, or even on the same information in a different order, then each of the neural networks will end up, after training, able to solve the task. However, the weight of particular synapses, or the analog of the receptive fields of particular neurons in the hidden layer, may be different in each of the networks. For each network, the particular combination of synaptic strengths it evolved to solve the problem differs, but the end result is the same.

The second characteristic appears to be more a bug than a feature. Backpropagation, unlike other characteristics of network models that are merely oversimplifications, seems at first glance to be very different from the way real nervous systems work, because it requires information—the error signal—to pass backward across synapses, from postsynaptic to presynaptic elements. The neural relevance of this requirement, as well as a more detailed description of the backpropagation process, is provided in part II.

THE NEUROBIOLOGY OF A BIOLOGICAL NEURAL NETWORK

The nature and diversity of biological nervous systems preclude any encapsulation of properties analogous to that provided above for artificial neural network models. In lieu of a general survey of nervous systems, this section presents two short overviews of a single biological preparation, the buccal ganglia of *Aplysia*, first from the perspective of neurobiology, and then viewed in terms of a neural network model. This dual perspective is appropriate because of similarities between the biological preparation and multilayer model networks. In particular, mechanisms for retrograde information flow can be postulated on the basis of the neurobiological data from this preparation and elsewhere, and used to examine classes of learning rules in model networks.

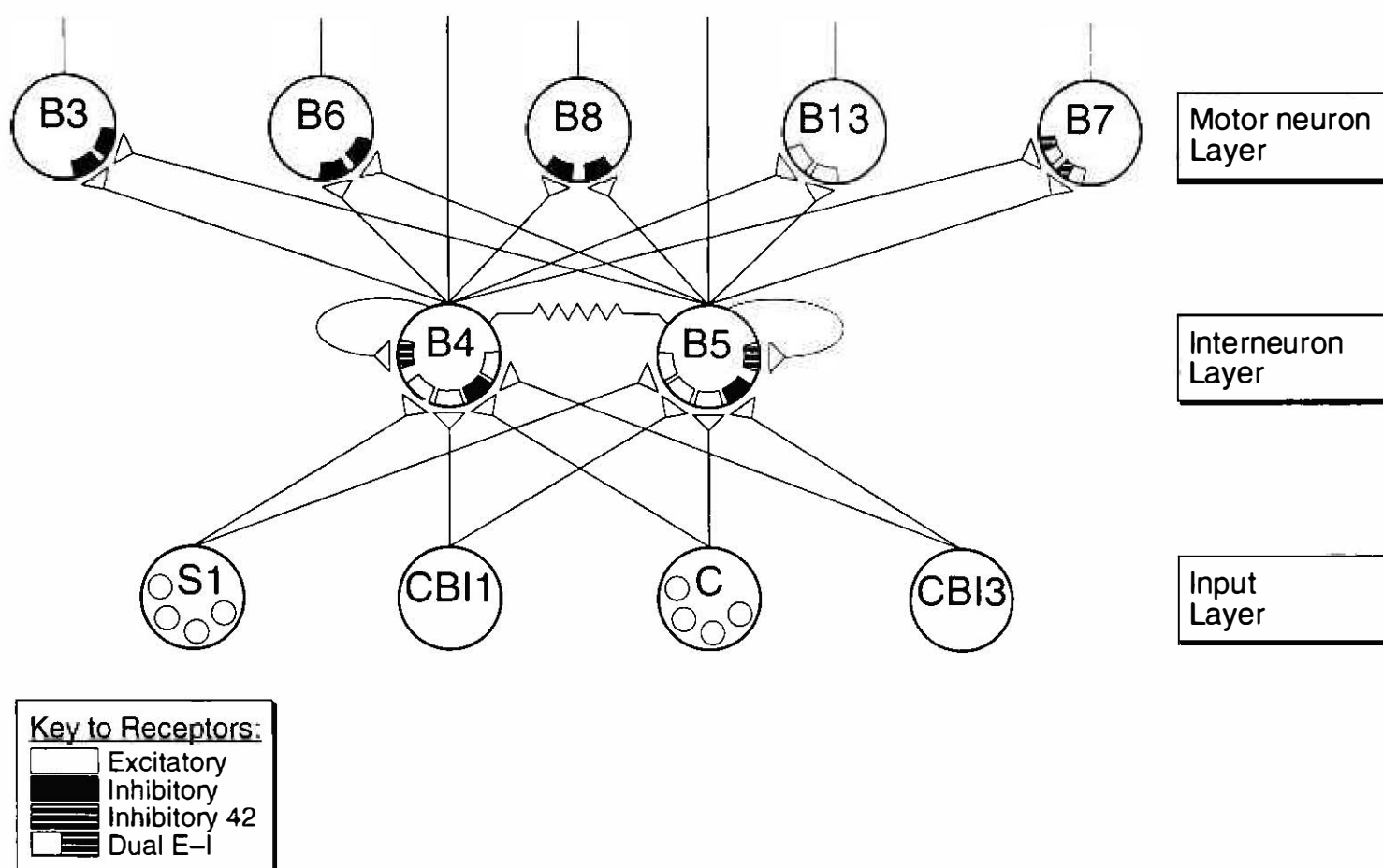
The Buccal Ganglia of *Aplysia*

The buccal ganglia, located near the head end of *Aplysia californica*, are a bilaterally symmetric cluster of neurons, several of which approach half a millimeter in diameter. More than 50 cells and several cell clusters in each ganglion, most with a symmetric mate in the other ganglion, have been identified (for reviews, see Gardner 1987; Kirk 1989). There are few direct interconnections between symmetric pairs of neurons; bilateral coordination usually reflects common synaptic input.

Within each buccal ganglion, a multilayer neural network is centered around two identified neurons, B4 and B5, referred to informally as interneurons. The neural circuits around B4 and B5 include most large cells in the ganglia. Each of the two interneurons monosynaptically innervates a common population of postsynaptic cells, producing in each follower a response qualitatively similar to that produced by the other interneuron. Each interneuron produces inhibitory postsynaptic potentials (IPSPs) in several identified neurons, including B3, B6, B8, B9, B10, B11, B15, and B16. Each also produces excitatory PSPs (EPSPs) in cell B13 and a diphasic excitatory-inhibitory PSP in cell B7 (Gardner 1971, 1977b; Gardner and Kandel 1972, 1977). In addition to this extensive innervation, each of the interneurons also inhibits itself (Gardner 1977a). Each of these connections, including the self-inhibition, appears to be monosynaptic and to be mediated by the neurotransmitter acetylcholine (ACh) (Baux and Tauc 1987; Gardner 1971; Gardner and Kandel 1977; Gardner et al. 1984; Ono and McCaman 1979). Synaptic potentials are large, with postsynaptic receptors electrotonically close to the soma, permitting low-noise voltage clamp analyses (Gardner 1980a; Gardner and Stevens 1980). The morphology of B4, B5, postsynaptic neurons, and relevant synapses has been reviewed by Kreiner et al. (1987).

The major function of the buccal ganglion circuit is the control of feeding. B4 and B5 not only are interneurons in the feeding circuit, but also serve sensory and motor roles as well. These cells receive input from several sources both within and without the buccal ganglia, thus serving as channels integrating multiple types of sensory input. From outside the ganglia, excitatory input to B4 and B5 is produced by cerebral ganglion mechanoreceptors (Rosen et al. 1982, 1991), as well as direct proprioceptive input (Jahan-Parwar et al. 1983); additional input arises from serotonergic neurons (Goldstein et al. 1984). These paired interneurons are also innervated from within the buccal ganglia by other identified cells and by the S cell cluster. S cells also excite one another and provide hyperpolarizing-depolarizing PSPs to B3, B6, B9, and B10 (Fiore and Meunier 1979).

Neurons B4 and B5 show both synchronous and asynchronous firing (Gardner 1977b). Several modes of activity of B4 and B5 are correlated with feeding behavior (Rosen et al. 1982; Susswein and Byrne 1988); the neuronal circuitry responsible for patterned activity in this system has been reviewed by Kirk (1989). B4 and B5 show patterned activity in at least two modes: loosely, a 20/min tonic firing and a 1/min phasic burst, tentatively identified with the behaviors of ingestion and rejection (Susswein and Byrne 1988). Several cells which are components of circuits involved in bursting activity of B4, B5, and their followers have been identified. These include interneuron CBI-2, which acts with B17 and B19 to drive a motor program, as well as a group including B31, B32, B33, B35, B37, B51, and B52, in which the interactions between membrane properties



Buccal Ganglion Neural Network

Figure 3.5 A simplified diagram of the biological neural network in each buccal ganglion of *Aplysia*, centered around identified neurons B4 and B5, which receive largely excitatory input from several neurons and clusters and in turn innervate many motor neurons. Identified neurons and neuron clusters are shown in layers to emphasize the imperfect analogy between this biological network and multilayer network models. Different postsynaptic responses are coded by schematic shading of receptors. For clarity, connections actually made between presynaptic and postsynaptic processes in the neuropil are shown schematically on cell bodies. The self-inhibitory synapses made by B4 and B5 are shown arising from recurrent collaterals; the weak electrotonic coupling between B4 and B5 is indicated by a resistor.

and connectivities are complex and not well understood (Plummer and Kirk 1990; Susswein and Byrne 1988; personal communications).

Firing of B4 and B5 inhibits motor neurons to buccal retractor muscles, excites some protractor motor neurons and mediates dual excitation-inhibition to others (Cohen et al. 1978; Lloyd et al. 1985). This feeding system demonstrates associative and nonassociative plasticity, pattern generation, and neuromodulation. Simultaneous optical recordings have revealed not only the large number of cells involved in eating behavior, but also many complex temporal relationships between different neurons (Morton et al. 1991).

In a sense, the *Aplysia* network provides as well an invertebrate analog to spinal networks of Ia afferents on homonymous α motor neurons, but with the added advantages of large quantal content, identified cells, low synaptic noise, and intracellular access to all neurons. Furthermore, in this system action potentials can be elicited or prevented by experimen-

tal manipulation, and short electrotonic distances reduce artifacts due to cable properties.

Similarities Between *Aplysia* Buccal Ganglia and Artificial Neural Networks

As the above summary of properties of the buccal ganglia shows, these cells, and the connections they make, can be used to derive some insight as to how small networks process information. The accessibility of individual neurons of the *Aplysia* buccal ganglia permits examination of these neurons and their interconnections in ways analogous to those in which an artificial neural network might be investigated. As an example, identified cells can be recorded from reliably and persistently, and several metrics related to synaptic strength at individual synapses can be analyzed (Gardner 1980b, 1990). Furthermore, the accessibility is such as to permit recording simultaneously from several cells, allowing multiple neurons and synapses to be characterized and examined in parallel, rather than sequentially. Finally, the architecture of the buccal ganglia can be viewed as if composed of multiple layers of similar elements. From this perspective, B4 and B5 represent a middle layer of a multi-layer network, receiving qualitatively similar information from several classes of similar cells, including the S1, S2, and extraganglionic clusters, and innervating in turn sets of similar higher-order neurons with synapses of varying strength. This similarity is emphasized by figure 3.5, which presents the wiring diagram of a subset of buccal ganglia neurons as if they were elements of a network. This diagram may also serve as a reminder that, just as in the case of models, what matters to a real network of neurons are the properties of the cells, the connectivity or wiring diagram, and the properties of the synapses that interconnect the cells, including the learning rules, if any, for their modification. Perhaps coincidentally, the self-inhibition shown for B4 and B5 parallels that employed by Frasconi et al. (1992), who designed a class of recurrent networks in which local feedback is restricted to the intermediate or hidden layer.

Again characteristic of network models, processing in the buccal ganglia is distributed among many cells, with widespread neuronal involvement in each of multiple behaviors (Morton et al. 1991). Such distributed processing is characteristic of other *Aplysia* ganglia (Zečević et al. 1989) and ganglia of other invertebrates (Lockery et al. 1989; see chapter 5), notwithstanding their restricted numbers of neurons.

Classes of neural network models have been criticized for implausibly transmitting output from a single unit through both positive and negative synaptic weights, and for permitting the sign of weights at single synapses to change (Shepherd 1990b). These properties are the neurobiological equivalents of individual neurons mediating both excitation and inhibition to different postsynaptic units without interposed inter-

neurons to change sign, and of individual synapses mediating both excitation and inhibition. However, each of these neurobiological analogs can be demonstrated in *Aplysia* buccal ganglia. Cells B4 and B5 not only mediate synaptic excitation and inhibition at different terminals, they produce conjoint excitation/inhibition in B7, in a time- and voltage-dependent manner. Extending the set of neuronal mechanisms to include this range of synaptic actions further strengthens the analogy between the buccal ganglia and network models. This behavior is additionally consistent with the *class two* neurocomputers described in chapter 2.

Synapses in buccal ganglia differ not only qualitatively, but quantitatively as well. As shown below, synaptic strengths of otherwise similar connections made by individual neurons can vary widely in a single preparation. This finding is again consistent with network models, in which synaptic weights between similar elements can be very different.

Finally, one additional characteristic of neural network models is that the result of training individual structurally identical examples of a single network plan on the same data can produce different combinations of synaptic strengths (for examples, see Ackley et al. 1985; chapter 7). Each of these individual instantiations can perform the information processing task, but the distribution of weights may be different, arising as a result of variations in the training set, or from the initially randomized weight values. Is there a biological equivalent to this property of network models? Given biological diversity within species, can one consider equivalent ganglia from individual animals as biological instantiations of a common network plan? Networks of identified cells in invertebrate ganglia provide a partial justification for this view. If individual members of a species are equated with instances of biological neural networks constructed on a common plan, their nervous systems would be expected to resemble one another in number of cells, qualitative properties of individual cells, and connectivity, but to differ in quantitative properties, chiefly synaptic efficacy. The next section presents evidence consistent with this view.

BUCCAL GANGLIA SYNAPTIC STRENGTHS DIFFER BOTH STATICALLY AND DYNAMICALLY

Biological Analogs of Synaptic Weight

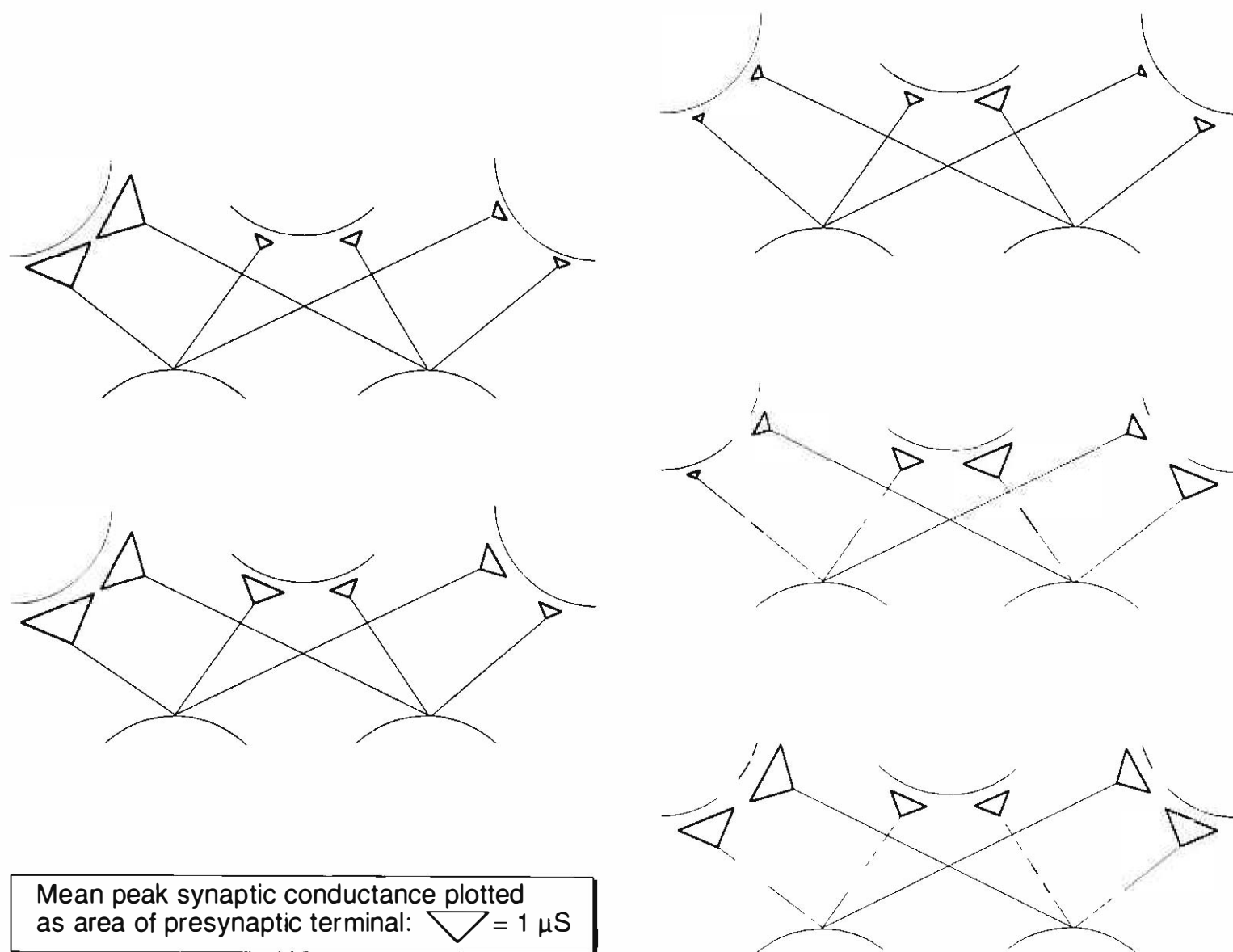
In previous work, I have calculated the time integral of conductance as a measure of synaptic efficacy (Gardner 1980b) and have proposed this as a biological analog to the synaptic weights characterizing artificial network models (Gardner 1989b). For the mechanistic analyses below, I consider peak synaptic conductance M and decay time constant τ , the two major components of time integral of conductance, separately. This analysis also reveals that variations in peak conductance are greater

than those in time constant, and that different synaptic strengths are more likely to reflect differences in peak conductance (Gardner 1986). In all cases, conductance rather than current is used as an analog of synaptic weight, eliminating contributions from the postsynaptic membrane potential or the synaptic mechanism's reversal potential. This definition also reduces possible dependence of biological equivalents to a particular w_{ji} on other values of w and other presynaptic activity. (The biological interpretation of w in this chapter is thus slightly distinct from that used in the chapter by Lockery and Sejnowski, who choose to incorporate not only synaptic conductance, but also electromotive driving force.) In this chapter, I further distinguish long-term, slowly varying processes that establish a set point characterizing the *static* synaptic strength from those producing short-term, phasic, *dynamic* variations in synaptic efficacy.

Static Differences in Synaptic Efficacy

The static strength of synapses between identified cells varies in two ways: between different cell pairs in the same ganglion, and at the same synapses between specific cell pairs in different animals. As an example, figure 3.6 plots peak synaptic conductance M of each of six synapses between B4 and B5 and three of their common postsynaptic neurons. These data, which are graphically indicated by scaling the area of the diagrammed presynaptic terminal, are presented for each of five ganglia, permitting comparisons both between synapses and between animals. Although B3, B6, B8, and their cousins are different cells, the synapses on them are indistinguishable by any criterion other than synaptic strength, and there is no evidence of any trend in synaptic strength separating one similar postsynaptic neuron from others. As most ganglia show a diverse mix of varying synaptic strengths, and no electrophysiological data suggest damage or injury to cells receiving larger smaller (or larger) postsynaptic currents, the differences do not appear to be artifacts of illness or poor condition of individual ganglia. Because these variations in synaptic strength are much larger than can be accounted for by such variables as temperature, they may reflect past (or continuing) influence of an adaptive variation. Part II of this chapter explores factors responsible for determination of static synaptic strength.

Variations in synaptic current amplitude exceed those in time constant. For the 56 synapses studied by Gardner (1989a), peak synaptic conductance M ranged from 0.017 to 1.35 μS , and τ from 12.5 to 31.0 ms (with few values above 27 ms). Time constant approximated a normal distribution much more closely than did either M or $\log_e M$; more extensive statistical analyses on a separate population of synapses is provided in Gardner (1990). In the population of 46 synapses studied in detail in part II, synaptic conductance amplitudes were distributed over a more than 70-fold range.



Synaptic Strengths Differ in the Buccal Ganglion Network

Figure 3.6 The differences in biological synaptic strength may be analogous to varying weight in trained network models, and to different instantiations of a single network type. Each of five diagrams indicates synaptic strengths at connections between B4 and B5, below, and three postsynaptic neurons, above. Each diagram presents data from one ganglion. Mean peak synaptic conductance, used as an index of synaptic strength, is plotted schematically as the area of the presynaptic terminal. Synaptic strengths for these data ranged from 1.60 to 0.04 μS .

Dynamic Variations in Synaptic Efficacy

Synaptic strengths vary dynamically as well as statically. Both components of synaptic strength of these synapses can be transiently altered by dynamic modulation (Gardner 1986). Simple manipulations of either the firing frequency or the steady-state soma membrane potential of the presynaptic neuron altered both postsynaptic conductance (PSC) amplitude and time course. For example, tetanic stimulation of B4 or B5 produced transient potentiation of both amplitude and time course of subsequent PSCs produced by the tetanized neuron. Increasing pre-synaptic firing frequency produced habituation of synaptic amplitude and time course, whereas conditioning DC depolarization of the pre-synaptic neuron enhanced PSC amplitude with variable effect on decay time constant.

Biological Synaptic Efficacy and Model Synaptic Weights

Synapses in the buccal ganglia network therefore resemble their engineered counterparts in network models in showing both dynamic modifiability and static differences in efficacy between similar elements (Gardner 1989a). Efficacy of *Aplysia* synapses is therefore a plausible biological analog to synaptic weights between elements of neural network models. Many questions concerning both specification and consequences of synaptic strength remain unanswered. For example, What is the relative importance to a biological neural network of absolute and relative synaptic strengths? Is the significance of a set of weights related only to the absolute ability to manipulate the membrane potential around threshold, or do relative weights possess significance for information processing?

Another functionally significant question asks if the wide variations observed in synaptic strength follow any pattern. One possibility is that although individual synaptic strengths differ, the aggregate synaptic efficacy of this network remains similar from animal to animal, with only the distribution of stronger and weaker synapses varying (Gardner 1990). However, summed synaptic strengths show variations similar to those seen for individual connections. For a set of 16 buccal ganglia, each with four synapses analyzed, the distribution of ganglion-average values of M varied uniformly over approximately an order of magnitude from less than $0.1 \mu\text{S}$ to $1 \mu\text{S}$. Part II presents evidence for a different pattern: that synaptic strengths do not vary randomly, but are constrained by individual postsynaptic neurons.

CELL AND NETWORK PROPERTIES OF THE BUCCAL GANGLIA TRANSCEND SIMPLIFYING ASSUMPTIONS OF NETWORK MODELS

In the buccal ganglion network, synaptic properties are more complex than those that usually characterize artificial neural networks. Some synapses made by B4 and B5 are excitatory, others inhibitory, whereas still others combine both excitatory and inhibitory actions in complex ways. Some synapses are sensitive not only to acetylcholine, the neurotransmitter released by B4 and B5, but also to choline, a product of hydrolysis of the neurotransmitter (Gardner et al. 1984). For some synapses, the presynaptic and postsynaptic elements are located on the same cell, so that a self-inhibitory autapse is formed (Gardner 1977a). Still other connections are electrotonic, effected by direct spread of current rather than mediated by a neurotransmitter (Gardner 1971).

In an earlier survey (Gardner 1987), the richness of cellular mechanisms and the number of disparate identified cells was used to emphasize the deficiencies of neural network elements. It was pointed out that properties such as dual-action synapses with conjoint excitatory-inhibitory actions,

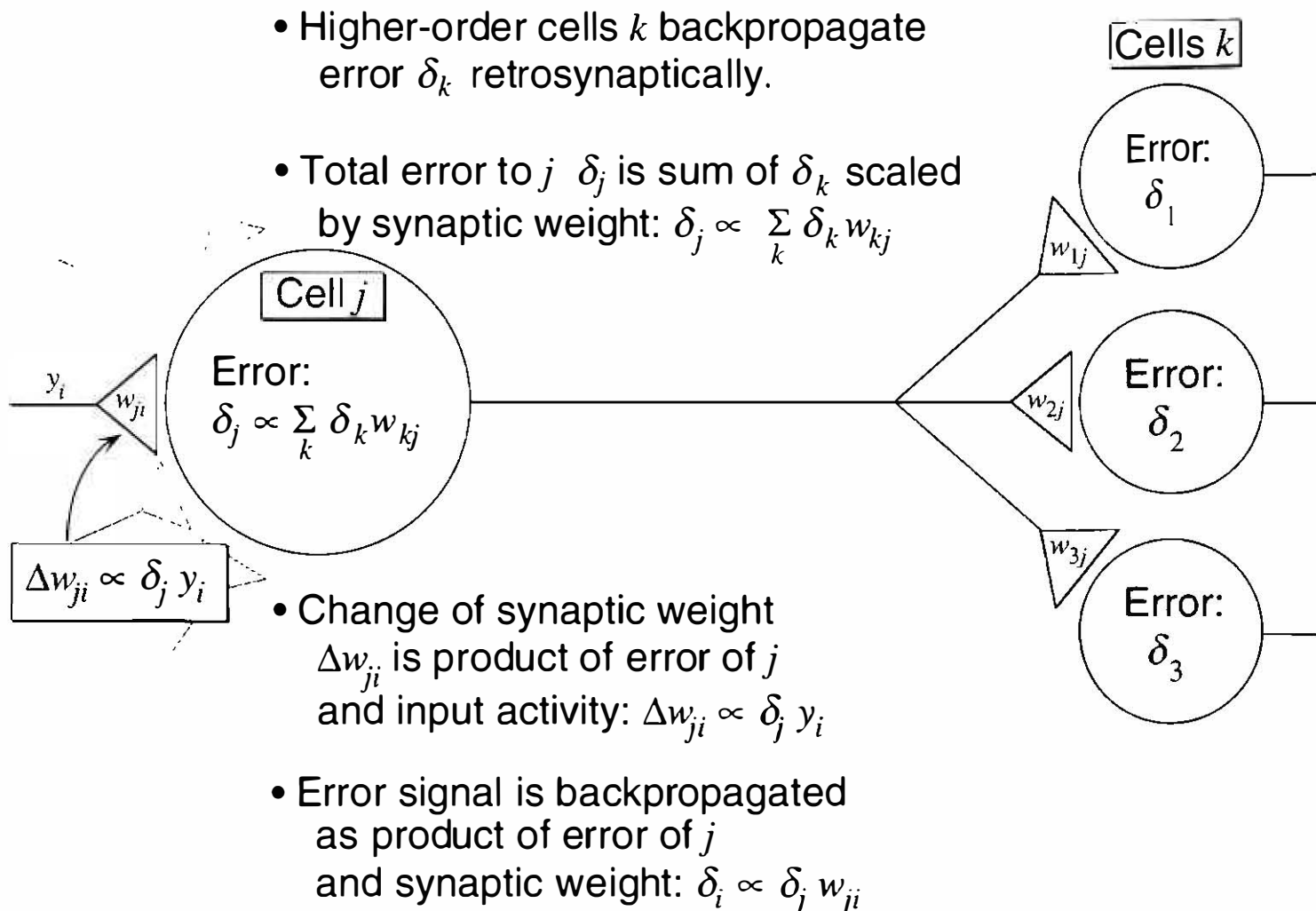
and synapses differing not only in amplitude but also in duration of action, were not common features of network models. Although composed of a limited number of cells, the buccal ganglia networks include an impressive repertoire of neuronal mechanisms, many of which transcend assumptions of standard neural network models. Functionally, these mechanisms include: cotransmitters, neuromodulation, both voltage-dependent and voltage-independent synapses, parallel channels for distributed processing, multi-action interneurons, synaptic potentials of varied amplitudes and durations, polyphasic synaptic potentials, differential desensitization of synaptic components, electrotonic junctions, sensitivity to breakdown products of neurotransmitters, and autapses.

Neurochemical diversity complements and in some cases underlies this functional diversity. In addition to ACh, neurotransmitters and co-transmitters utilized by buccal neurons include serotonin, dopamine, small cardioactive peptides (SCP), cholecystokinin (CCK), egg-laying hormone (ELH), the myomodulin family, FMRFamide, and buccalin, as well as other peptides (Cropper et al. 1987; Gerschenfeld and Paupardin-Tritsch 1974; Goldstein et al. 1984; Ono 1989; Schwartz and Shkolnik 1981; Sossin et al. 1987; Whim and Lloyd 1989); second messengers and target channels are identified for many of these modulatory actions.

II. Biological Networks and Reverse Information Transmission

One aspect contrasting biological and artificial neural networks that was covered only implicitly in part I is the way in which each specifies and modifies the strengths of synapses. This part presents in detail evidence that synaptic strengths in *Aplysia* buccal ganglia are partially specified by the postsynaptic neuron. I further show that these differences, although specified postsynaptically, represent differences in quantal transmitter release. Several classes of mechanisms that could be responsible for setting and altering synaptic strength are reviewed, each of which requires some information to flow retrosynaptically, from postsynaptic to presynaptic neuron. Evidence for similar retrograde information transmission in other preparations is reviewed. The goal is not to claim that all synapses, or even all *Aplysia* synapses, alter quantal release by postsynaptic or heterosynaptic modulation, but to show that this is a plausible class of mechanisms.

The relevance of these data to neural network models is presented as introduction and conclusion to the experimental findings. I begin this part with an examination of the use of retrograde information flow to set synaptic weights in backpropagation models, and conclude it with an attempt to relate these network learning rules to possible biological mechanisms.



Network Equations for Backpropagation

Figure 3.7 Neural network description of backpropagation in a model neuron in the j th layer, receiving input from the i th layer and innervating in turn cells in the k th layer. Error signals δ characterizing k higher order cells (three shown) and postsynaptic cell j are indicated within the schematic cell bodies; synaptic weights w are indicated within triangles representing presynaptic endings. Note that the error signal δ is used both to backpropagate the error to presynaptic cells and to adjust synaptic weights; the first case depends on the product of δ and w , and the second on the product of δ and y , where y is presynaptic activity. Notation follows Rumelhart et al. (1986a), except that δ is conventionally used to represent the error signal. For simplicity, the proportionality operator $\propto$ is intended to incorporate the derivative of the nonlinear scaling function f . Compare to figure 3.8.

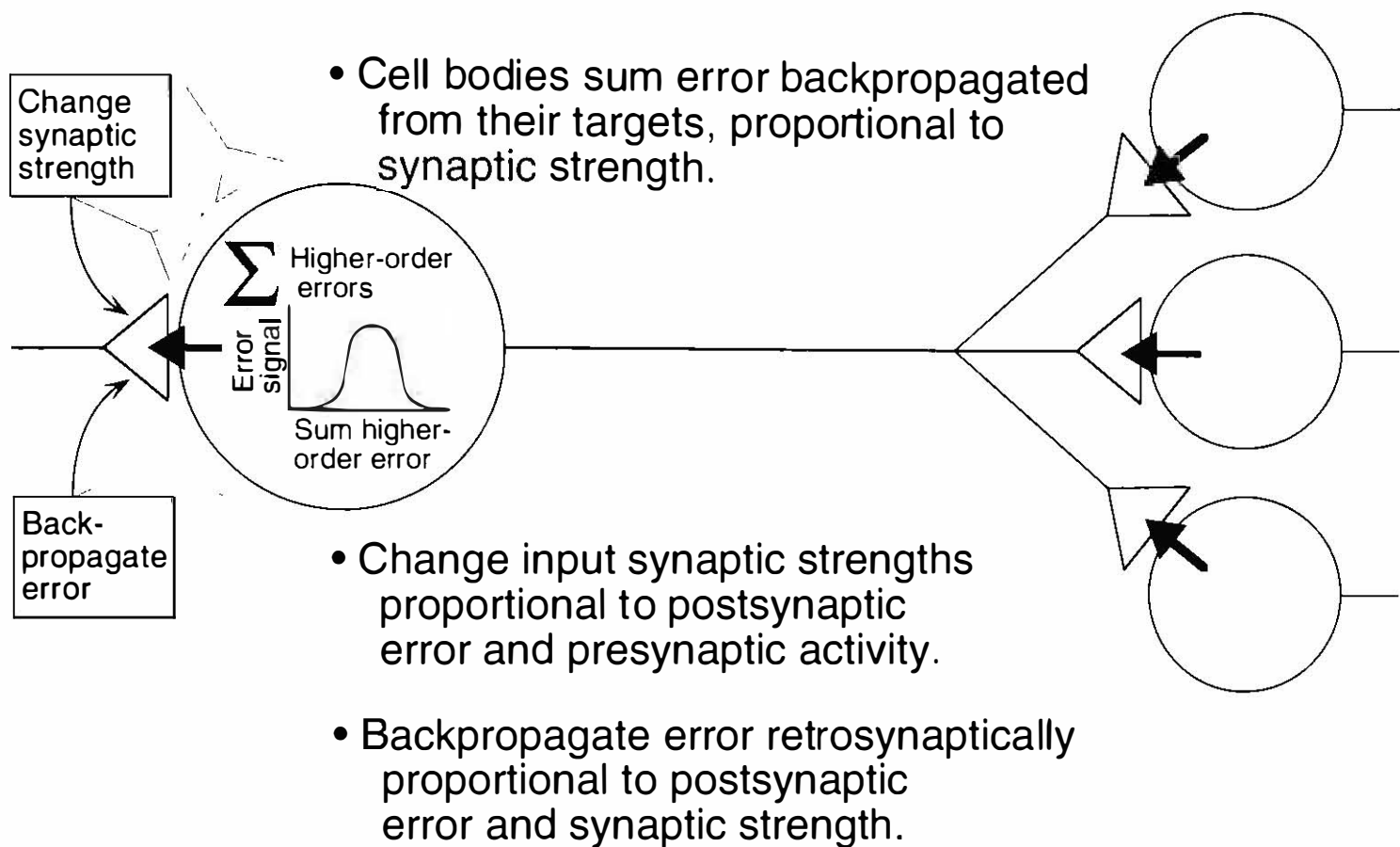
BACKPROPAGATION IMPOSES A REQUIREMENT FOR RETROGRADE INFORMATION TRANSFER

Backpropagation and Error Minimization in Network Models

The presentation of multilayer network models in part I did not explicitly consider the algorithms used for changing synaptic weights so as to minimize error and optimize the function of the neural network. An analysis of backpropagation, a form of gradient descent, is prerequisite to examination of its neurobiological relevance. Additional aspects of gradient descent are reviewed in chapter 5.

The rules of backpropagation are presented in each of two ways, just as model elements were in part I. In this case, the defining equations are discussed first, followed by their neurobiological equivalents.

- Higher-order cells backpropagate error retrosynaptically.



Neurobiology of Backpropagation

Figure 3.8 Neurobiological description of backpropagation in a hypothetical neuron receiving conventional synaptic input and innervating in turn three other neurons. Error signals (indicated by broad arrow) from three higher-order cells are transmitted retrosynaptically to innervating terminals. Cell bodies sum weighted higher-order error signals from their terminals and scale, for example, by the indicated function. This scaled error signal is used to adjust the synaptic strengths of active presynaptic afferents. Note that the error signal is used both to backpropagate the error to presynaptic cells and to adjust synaptic weights; the first case depends on the strength of the synapses and the second on presynaptic activity. Compare to figure 3.7.

Figures 3.7 and 3.8 provide a description of the process on the cellular level, using the same prototypic model neuron presented in figures 3.1 and 3.2. Because backpropagation is a retrosynaptic process, the direction of information flow in figures 3.7 and 3.8 is right to left. Using an external teacher, error signals are generated in each output unit, as scaled differences between desired and calculated output values. For cells in all other layers, such as cell j of figure 3.7, the error signal is computed from the errors characterizing higher-order units. Each of the k higher-order cells that j innervates is characterized by error δ_k for each input pattern. The total error of j is a function of the scaled sum of the δ_k , each scaled by the weight of the appropriate synapse w_{kj} :

$$\delta_j \propto \sum_k \delta_k w_{kj}. \quad (3.3)$$

This error δ_j is in turn backpropagated to each of the cells i that innervate

j . The error signal characterizing each cell not only contributes to the error of lower-order cells, it is used to change weights at each level so as to minimize total error. The alteration in weight of the synapse made by cell i on cell j , Δw_{ji} , is proportional to the product of δ_j and presynaptic activity y_i :

$$\Delta w_{ji} \propto \delta_j y_i \quad (3.4)$$

Figure 3.8 presents the same process from a more neurobiological perspective. Higher-order cells on the right are each shown transmitting an error signal, perhaps as a retrograde chemical transmitter substance, to incident presynaptic terminals. Each of these presynaptic terminals receives this retrosynaptic messenger, scales the error signal received proportional to its anterograde synaptic strength, and conveys it back along axonal processes to the presynaptic cell body. The cell body sums these weighted higher-order errors and scales the result by the nonmonotonic function shown. This resulting error signal is then propagated backward to still lower-order cells. In addition of this propagation and recalculation of the error signal itself, the error transmitted by each postsynaptic cell is used to adjust the strengths of the synapses innervating it. This alteration of synaptic weight is proportional to the postsynaptic error δ_j , as before, but also to the presynaptic *activity* y_i , rather than to the synaptic strength.

In sum, there are six relevant points that characterize backpropagation: (1) An error signal is transmitted retrosynaptically, from postsynaptic to presynaptic neuron. (2) The retrosynaptic signal characterizes the entire postsynaptic neuron, rather than a partial region of the cell. (3) The retrosynaptic signal is received by individual presynaptic terminals. (4) The retrosynaptic signal, as scaled by the *weight* of the appropriate synapse, is utilized to contribute to the error signal of the presynaptic cell. (5) The retrosynaptic signal, as scaled by the *activity* of the presynaptic cell, is utilized to alter the weight of the appropriate synapse. (6) The purpose of retrosynaptic signaling is the adjustment of synaptic strength, in contrast to the purpose of conventional synaptic transmission, which is the ongoing processing of information from presynaptic to postsynaptic neuron.

Backpropagation and Neurobiology

Such learning rules as backpropagation seem fundamentally different from the way real networks work. The requirement for reverse transmission of information across synapses in order to perform error correction is a distinctly nonbiological characteristic of neural network models. Chemical synaptic transmission utilizes distinct mechanisms in presynaptic and postsynaptic structures that are specialized, respectively, for transmission and reception of neurotransmitter molecules. The classical view of neuronal transmission suggests that information

should pass only from the presynaptic to the postsynaptic cell, and not the other way.

Is it therefore necessary to exclude classes of algorithms employing backpropagation from consideration by the neurobiologist? This chapter, without in any way suggesting the existence of backpropagation or similar mechanisms in any biological nervous system, presents evidence from disparate biological preparations consistent with the existence of retrograde information transmission across chemical synapses. Moreover, this biological retrosynaptic signaling appears to play a role in the modulation of synaptic efficacy, analogous to the role of backpropagation in network models. This evidence suggests that models incorporating reverse information flow need not be rejected solely on the basis of this feature.

Backpropagation is not the only successful algorithm for error minimization of multilayer network models, nor is it the only one that imposes the apparently non-neuromorphic requirement for retrograde information transfer. A useful alternative algorithm exists, whose reliance on gradual reduction of a state variable that can redistribute weights according to a popular distribution has given it the name Boltzmann machine (Ackley et al. 1985). The architecture of these networks depends on symmetric synapses, so that each neuron pair is interconnected bidirectionally. The requirement for symmetric weights is derived from spin-glass-like energy minimization and permits single units to act locally so as to reduce global error. Even if these synapses did not explicitly introduce the equivalent of retrograde information flow by forcing symmetry, the algorithm for weight change incorporates the joint activity of presynaptic and postsynaptic units. This implicit retrosynaptic or transsynaptic requirement is thus the same as that specified by Hebbian mechanisms.

In classical backpropagation (Rumelhart et al. 1986a), computation of the error signal utilizes input from an external teacher, another feature that has been criticized as nonbiological. However, backpropagation does not arise solely from the utilization of a teacher. Teacherless algorithms can also learn effectively by means of backpropagation; indeed, removal of reliance on the external teacher may allow more parallel, more local, and more modular learning (Becker and Hinton 1992).

The apparent inconsistency between the requirements imposed by the backpropagation model and verified neuronal mechanisms has been addressed in the past in a number of ways. Neural network modelers are often careful to point out that the elemental "neurons" and "synapses" of their model need not correspond exactly to their neurobiological equivalents, either in computational ability or in mechanism (Rumelhart et al. 1986b). The discrepancy can be elided in several ways. One is to assume that each "neuron" corresponds to an assembly, small circuit, or column of biological neurons. However, as Shepherd (1990b) and others have pointed out, the regional complexity and heterogeneity of

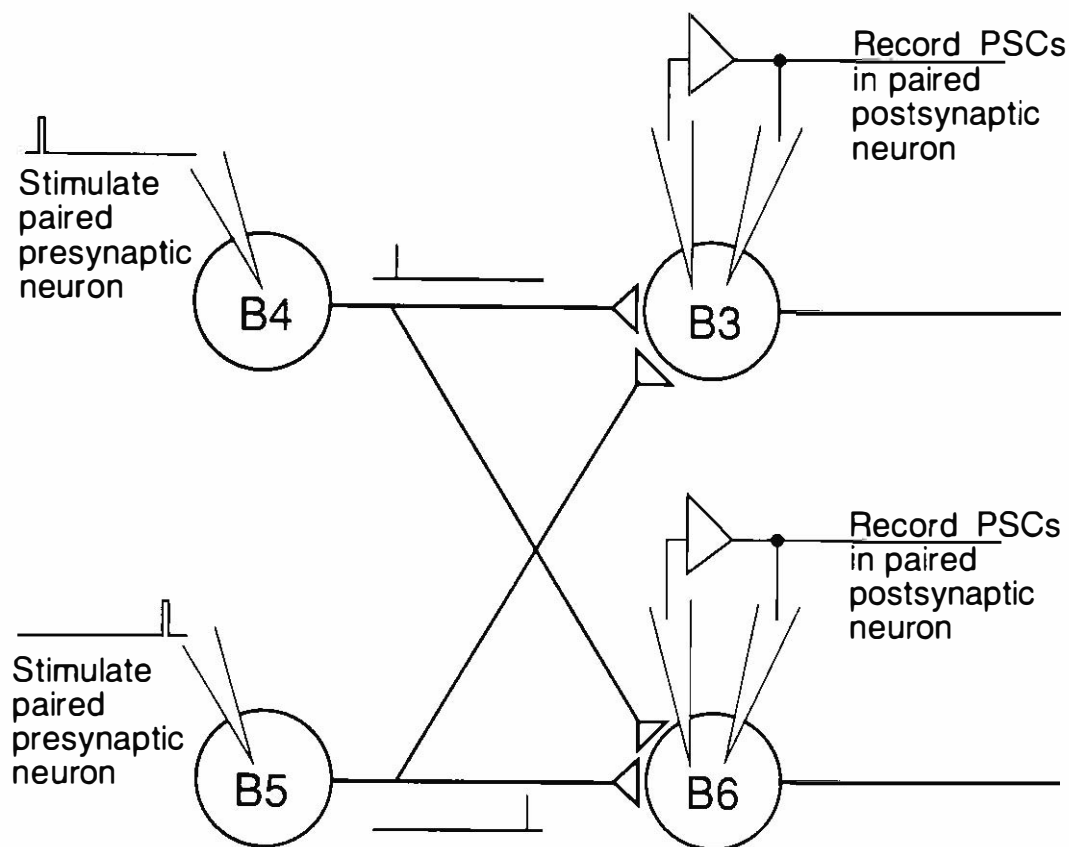
single neurons can endow them with a computational power greater than that of single model elements. Another response is to claim similarity for the solutions in weight space achieved by nervous systems and by models presented with analogous tasks, while distinguishing between the methods used by the two systems. In chapter 5, Lockery and Sejnowski stress that backpropagation is appropriately regarded as a scheme for deriving near-optimal solutions—solutions that are independent of the methods used to derive them—rather than as a model for a neurobiological process. In this view, backpropagation networks are “shallow” rather than “deep” in that they model the solution to an information processing task, rather than the algorithm responsible for the solution.

Given the similarities between biological and model networks sketched in part I, do neurobiological mechanisms exist that could mediate retrosynaptic information transfer, and so provide some neuromorphic justification for backpropagation and similar models? In a recent review of these issues, Zipser and Rumelhart (1990) have proposed dual pathways to implement backpropagation, separate paired forward and backward networks with distinct but interconnected elements. This approach has the dual advantages of accounting for parallel afferent/efferent pathways seen at many levels of the nervous system, without requiring any overt retrosynaptic mechanisms. Indeed, their justification includes the statement:

The concept of sending some kind of signal very rapidly back down an axon from postsynaptic to presynaptic neuron is inconsistent with any reasonable interpretation of current neurophysiological knowledge.

Is this statement unduly restrictive? While not treating the requirement for retrograde axonal transport of the signal, it is possible to address the apparently more stringent rejection of neurobiological mechanisms capable of mediating backward transmission across synapses. Evidence from two very different preparations argues that rejection of retrosynaptic mechanisms for the specification of synaptic strength may be less well grounded than previously thought. Although backpropagation is unlikely to be utilized by biological nervous systems, classes of mechanisms incorporating retrograde synaptic information flow are not necessarily inconsistent with neurobiology.

First, I use data recorded from the *Aplysia* buccal ganglion network described in part I to derive indirect evidence for a retrosynaptic contribution to specification of synaptic strength. Second, recent evidence from mammalian hippocampus implicating a retrosynaptic factor in the induction of one form of synaptic plasticity is combined with complementary *Aplysia* data to suggest possible mechanisms for this biological analog to backpropagated weight change.



Recording From a Biological Neural Network

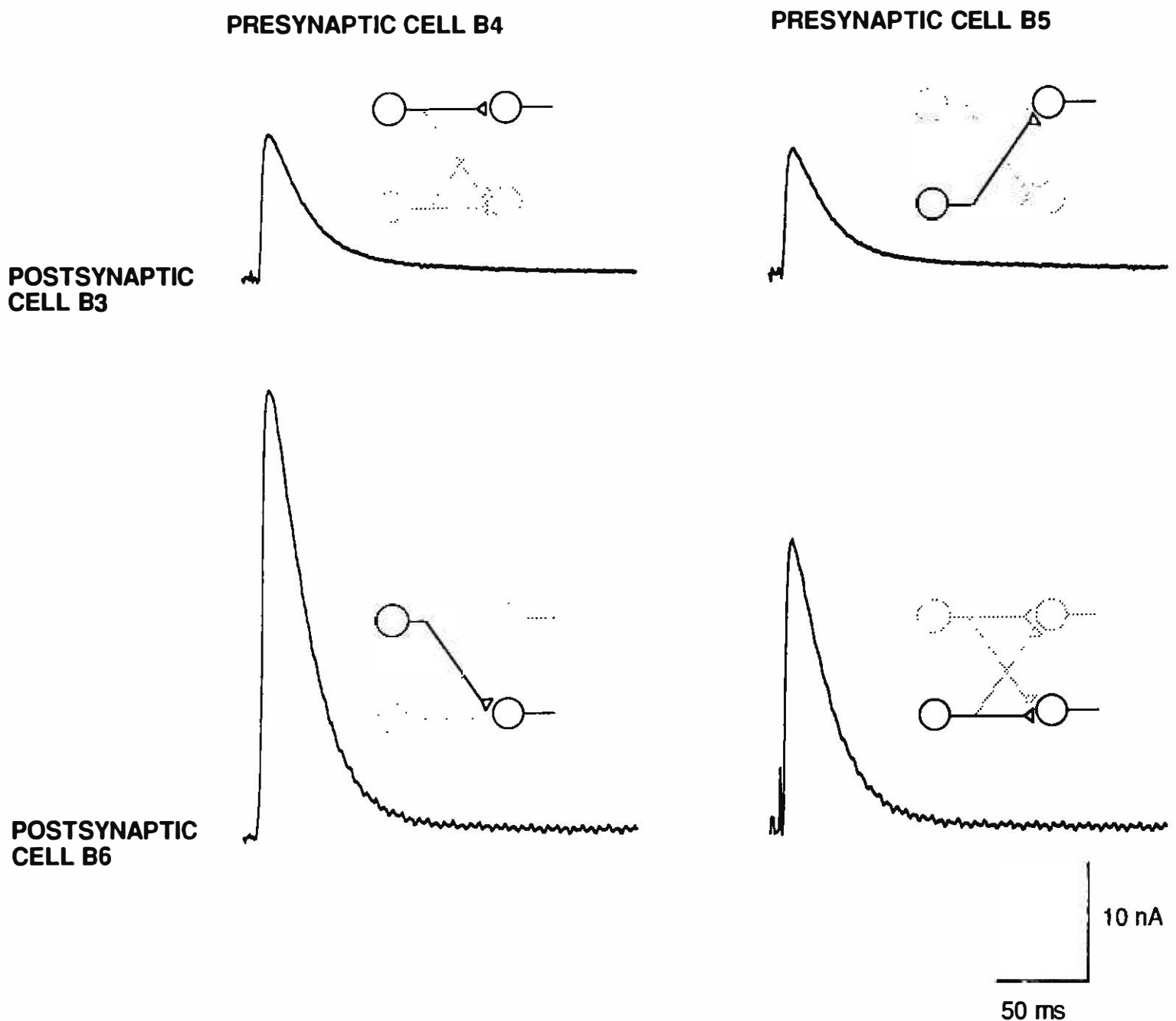
Figure 3.9 Schematic diagram of experimental protocol for comparing PSCs produced in similar postsynaptic cells by paired presynaptic neurons. Cells B4 and B5 are stimulated individually and alternately by means of short depolarizing pulses delivered through single intracellular electrodes. Resulting action potentials are indicated schematically. Two postsynaptic cells innervated by B4 and B5 (B3 and B6 shown here) are each voltage-clamped with two electrodes, permitting recording of simultaneously produced PSCs and subsequent comparison. (Adapted from Gardner 1989a, with permission of the American Physiological Society.)

BUCCAL GANGLIA SYNAPTIC STRENGTHS ARE SPECIFIED BY POSTSYNAPTIC NEURONS

Recording and Pairing of Buccal Ganglia Synaptic Currents

Analyses of synaptic strength in the buccal ganglia depend on network properties. Recall that two nearly identical presynaptic cells innervate at least eight similar motor neurons, and it is the existence of pairs of synapses in this circuit that permits examination of synaptic strength. As described below, using these paired synapses to analyze underlying quantal parameters also permits separating sites of specification of synaptic strength from the sites containing the mechanisms, thereby providing indirect evidence for retrosynaptic information flow.

These analyses depend on simultaneous recordings from a four-cell network, consisting of paired presynaptic neurons B4 and B5, and a pair of common postsynaptic cells (figure 3.9). This network contains four qualitatively similar inhibitory synapses, whose properties are briefly sketched in part I and described in greater detail by Gardner and Stevens (1980) and Gardner (1980a, b). Uncontaminated recording of postsynaptic currents from each of these four synapses requires independent



Individual Postsynaptic Currents Reflect Differing Synaptic Strengths

Figure 3.10 The four traces are examples of single unaveraged postsynaptic currents recorded at four synapses in the same preparation. (Left) Currents recorded simultaneously in postsynaptic cells B3 and B6 in response to stimulation of presynaptic cell B4. (Right) Currents recorded simultaneously in postsynaptic cells B3 and B6 in response to stimulation of presynaptic cell B5. The upward-going traces represent increases in outward current in cells clamped 25 mV more positive than the synaptic reversal potential. Both postsynaptic cells are recorded at the same gain; the larger noise observed in the B6 trace reflects a greater sensitivity of one voltage clamp to interference radiated by the PDP-11 computer. (Reprinted from Gardner 1990, with permission of the American Physiological Society.)

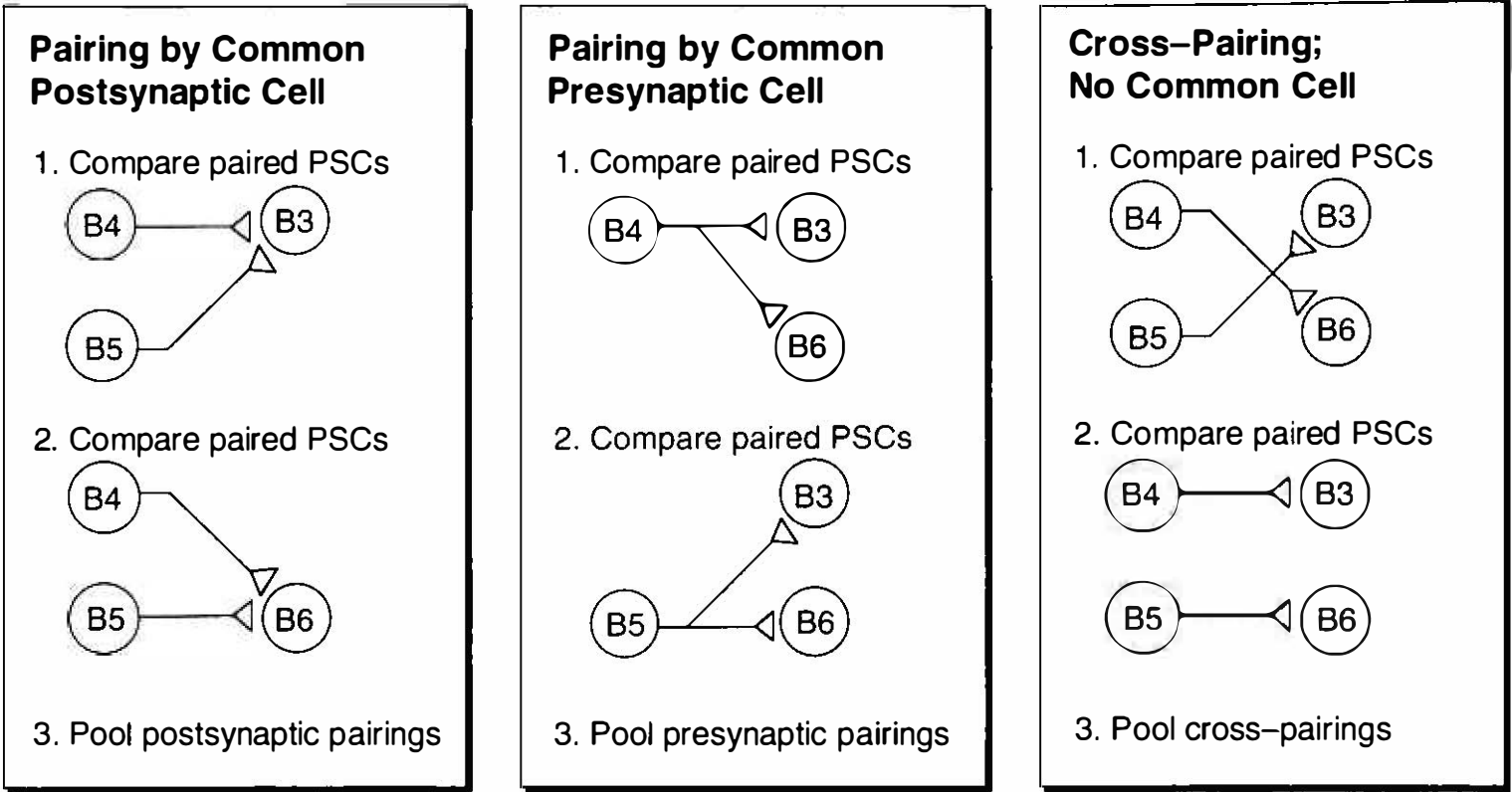
stimulation of each presynaptic cell by an individual intracellular microelectrode, and penetration of each postsynaptic neuron by two intracellular electrodes for voltage clamping. The methodology is described in greater detail in Gardner (1990).

To extract neurobiological parameters corresponding to synaptic strengths from each of the synapses of this network, presynaptic cells B4 and B5 are stimulated individually and alternately. Each presynaptic action potential produces an inhibitory postsynaptic current in both postsynaptic neurons. Examples of individual unaveraged postsynaptic currents produced in postsynaptic cells B3 and B6 by sequential stimulation of B4 and B5 are shown in figure 3.10. The two synaptic currents on the left were produced simultaneously in postsynaptic cells B3 and B6 following a single action potential in B4; the currents on the right are produced simultaneously in the same postsynaptic pair following an action potential in B5. As previously shown in figure 3.6, these PSCs differ in amplitude, reflecting different static strengths at each synapse.

Amplitude and time course of individual synaptic currents reflect the influence of both static and dynamic processes that vary synaptic efficacy. In order to isolate static determinants of synaptic strength, synaptic currents are recorded repetitively for several minutes to hours. Each of the four synapses of the network can then be characterized by mean values of synaptic amplitude and duration that lessen the contribution of dynamic factors or of noise. A fundamental assumption underlying this analysis is that these values reflect a past history of influence by plastic mechanisms. An alternative explanation, that the strengths of synapses are determined by genetic or fixed developmental factors, is inconsistent with the observation that strengths of identified synapses vary widely from one preparation to another.

This experimental procedure thus yields values of mean peak amplitude M and decay time constant τ characteristic of each of the four synapses in the four-cell network. From the perspective of each one of the four synapses in the network, the three other synapses can be uniquely characterized as sharing a common presynaptic neuron, sharing a common postsynaptic neuron, or having neither presynaptic nor postsynaptic neuron in common. As a result, *pairs* of synaptic parameters can be grouped in each of three ways, shown in figure 3.11. As a result of the synaptic convergence and divergence characteristic of the network, pairing the same data in each of several ways—by common postsynaptic or presynaptic cells, or by cross-pairing—is internally controlled, with each value of M or τ appearing in each of the pairings.

The rationale underlying these pairings is that similarities between paired values observed with one of the pairing schemes but not the others reflect a *specification*, or limitation of that parameter by the common synaptic element. For example, similarities seen in the common postsynaptic neuron case, pairing values produced by convergent terminals of B4 and B5 synapsing on a single postsynaptic cell, reflect



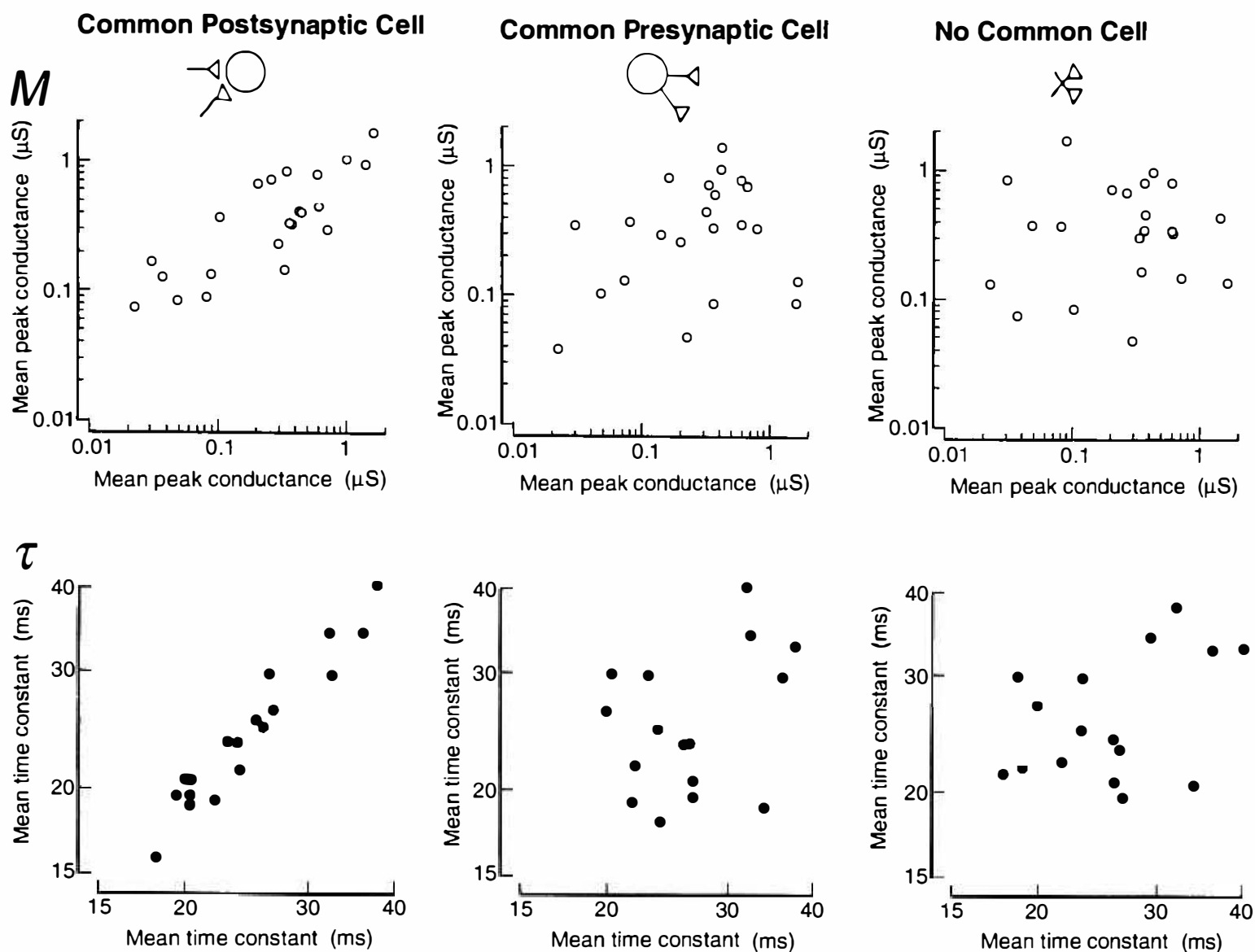
Pairings Test Presynaptic or Postsynaptic Similarity

Figure 3.11 Schematic diagrams of protocols for comparing parameters characterizing pairs of synapses in a network. Parameters describing each of the four synapses recorded by the method of figure 3.9 are paired in each of the ways shown in the three panels. This procedure provides internal controls, as each data point is incorporated in each protocol, differing only in the pairing. (Left) The common postsynaptic cell pairing compares data from synapses made by terminals of B4 and B5 converging on a single postsynaptic neuron. Convergent pairings on each follower cell are then pooled. (Center) The common presynaptic cell pairing compares data from synapses made by divergent terminals of individual presynaptic neurons on different postsynaptic cells. Divergent pairings from B4 are then pooled with those from B5. (Right) The cross-paired case compares two synapses sharing neither common postsynaptic nor common presynaptic cells. The two cross-pairings are then pooled. Although postsynaptic cells are shown as B3 and B6, data were obtained from additional postsynaptic cells as well.

postsynaptic specification of the synaptic parameter. In contrast, similarities seen when terminals of a single presynaptic neuron, such as B4, diverge to innervate separate postsynaptic neurons would reflect a presynaptic specification of the synaptic parameter. Similarities included in the third pairing case, in which the compared synapses share neither a presynaptic nor a postsynaptic neuron, might be ascribed to general properties characterizing the buccal ganglion as a whole, or the animal from which it was obtained (Gardner 1990).

Postsynaptic Specification of Synaptic Amplitude and Time Course

The six pairs of values of synaptic peak amplitude M and decay time constant τ characterizing each pairing of a synaptic parameter in each network are plotted in figure 3.12. The layout of the three vertical columns of the figure corresponds to the pairing schemes of each of the



Synaptic Strength is Postsynaptically Specified

Figure 3.12 Biological synaptic strength in the buccal ganglion network is specified by postsynaptic neurons. Each of the two major components of synaptic strength is compared at paired synapses; mean peak conductance M on the top row and PSC decay time constant τ on the bottom row. The same data are presented in each of the three graphs in each row, which differ only in the pairing scheme. The arrangement of each row corresponds to the three protocol panels of figure 3.11. (Top) Each data point compares M for paired sets of synapses, plotting one value on the horizontal axis against the paired value on the vertical axis. (Top left) Values of M paired by common postsynaptic cell are significantly more similar than randomly paired ($P = .0001$). (Top center) Pairing the same data by common presynaptic cell reveals no similarity ($P = .60$). (Top right) Cross-pairing M again reveals no similarity ($P = .79$). ANOVA for $\log_e M$ rather than M yields similar values. (Bottom) Each data point compares τ for paired sets of synapses, plotting one value on the horizontal axis against the paired value on the vertical axis. (Bottom left) Values of τ paired by common postsynaptic cell are significantly more similar than randomly paired ($P = .0001$), or to either of the alternate pairing schemes. (Bottom center) Pairing the same data by common presynaptic cell also reveals some similarity ($P = .05$), but no more than is shown by cross-paired data (bottom right, $P = .02$), suggesting that factors global to the ganglion or specific to experimental conditions may account for similarity in both of these latter cases. Although the same synapses are being compared in both top and bottom rows, exclusion of nonexponential and noisy decay reduced the number of data points in the τ comparisons.

three panels of figure 3.11. Graphs in the top row compare M for each of the three pairing schemes; the bottom row correspondingly compares τ . Each graph plots pairings; within each graph, the horizontal and vertical coordinates of each point plot a synaptic parameter at paired synapses. The same data are presented in each of the graphs in a row, differing only in the particular pairing scheme.

The top row of figure 3.12 shows that peak amplitude of synaptic currents is specified by the postsynaptic neuron. Pairings in the common postsynaptic neuron case, on the top left, show paired synapses to be similar to one another, with values arrayed near the diagonal representing equivalence. When the same data are paired by common presynaptic cell, or cross-paired without common cell, as shown in the top center and top right graphs, little or no similarity is seen. Values of M are thus similar for synapses formed by inputs from different presynaptic cells converging on a common postsynaptic cell. In contrast, values of M are different for synapses formed by divergent terminals of a single presynaptic neuron on different postsynaptic cells.

Analysis of variance was used to confirm the apparent similarity seen in the common postsynaptic neuron case. Peak conductances of PSCs produced in the same neuron by different, but paired, inputs are significantly more similar than unpaired ($P < .0001$), whether M or $\log_e M$ is used for the comparison. On the other hand, no significant relation was seen for presynaptic pairing.

Similar analyses for τ , rather than M , are shown on the lower set of graphs in figure 3.12. Again, these data show that PSC duration, as well as amplitude, is determined by the postsynaptic cell. When paired by postsynaptic cell, as shown in the lower left, time constant values are highly similar ($P = .0001$). Possibly reflecting properties of individual buccal ganglia or animals, both presynaptic pairing and control cross-pairing show lesser, but significant, similarity.

These results have utilized pairing to investigate the site responsible for long-term specification of synaptic strength in a neural network. Synaptic strengths are similar for inputs converging on a cell but different for branches diverging onto different targets, suggesting that they are specified by the postsynaptic neuron (Gardner 1989a, 1990). The findings imply not only that postsynaptic neurons specify static values of both peak amplitude and time course of decay, but also that this specification may act cell-wide to affect convergent inputs on that neuron. Apparent postsynaptic specification of synaptic efficacy could in some protocols be an artifact of such nonsynaptic characteristics of individual postsynaptic neurons as membrane potential or input resistance. Because all postsynaptic currents in these experiments were recorded in voltage-clamped postsynaptic neurons at controlled potential differences from synaptic reversal, such characteristics are controlled for, and do not contribute to the similarities of synaptic efficacy observed.

Independence of Synaptic Strength in a Biological Network

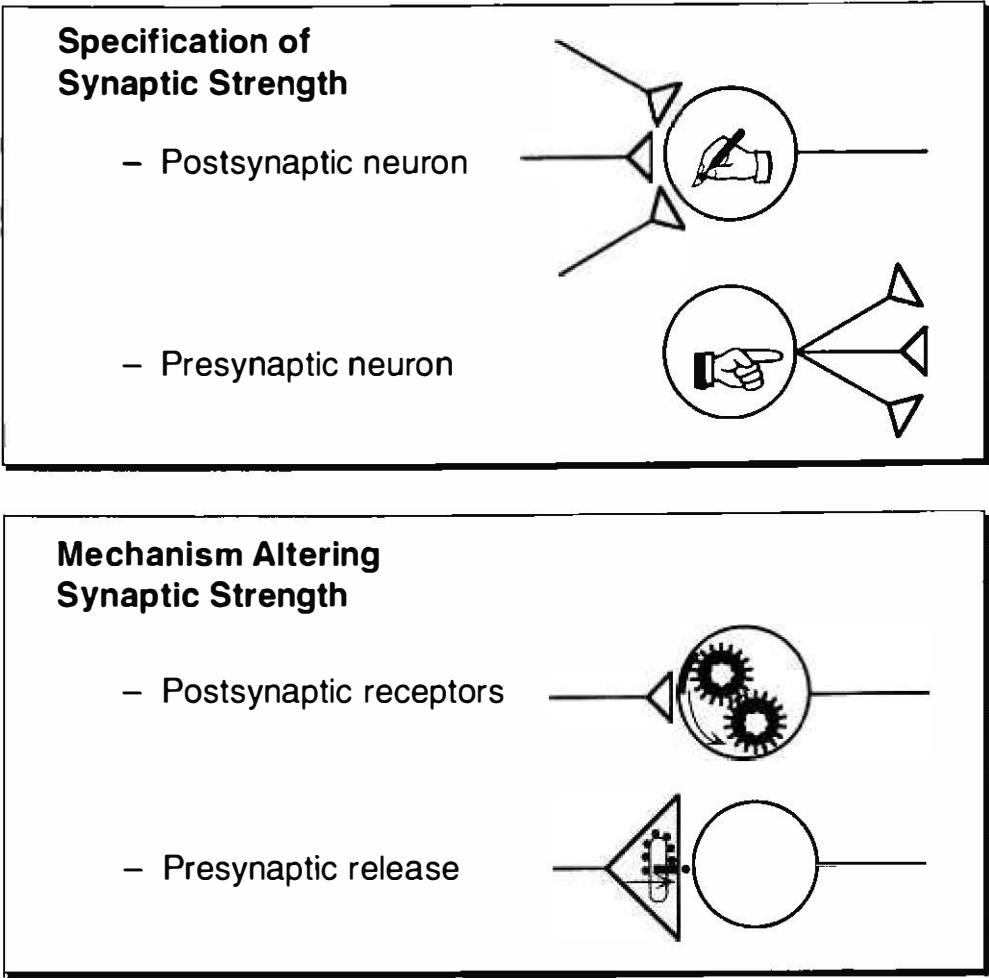
Dissimilarities in synaptic strength seen in similar postsynaptic neurons innervated by divergent terminals of a single presynaptic neuron suggests that the neural message communicated is different for different recipients. These quantitative differences complement the differences of synaptic sign and action already observed from individual presynaptic neurons. Both findings are similarly consistent with neural network models, in which both positive and negative synaptic weights are individually modifiable (Gardner 1989a).

The similarity of synaptic strength on single postsynaptic cells highlights the role of paired interneurons B4 and B5 in distributed processing. Even though outputs from a single interneuron may differ in sign of synaptic action as well as efficacy on different postsynaptic neurons, individual postsynaptic neurons tend to weigh input from both presynaptic neurons similarly. This observation suggests that one role of paired neurons B4 and B5 is as a partially parallel channel for distributed information transmission to their common postsynaptic population. However, the two neurons are not merely redundant. Paired convergent outputs may be similarly specified, but they are not identical, and B4 and B5 show differences in excitability and several patterns of both independent and common activity.

Synaptic Strength May Be Specified at a Site Different from the Biophysical Mechanism

Many of the processes of synaptic transmission may be separated into two phases, depending on location. In the presynaptic neuron, mechanisms are specialized to release a number of prepackaged quanta of neurotransmitter, whereas the postsynaptic neuron makes available a certain number of receptor or receptor/channel molecules and the processes they influence.

It is necessary to distinguish between the site of *specification* of synaptic strength, determined above by similarity of pairings, and the site of the *mechanism* that is altered in order to modify synaptic strength (figure 3.13). Identification of the postsynaptic cell as the site of specification does not identify the mechanism that *alters* synaptic strength, nor does it localize that mechanism to a presynaptic, a postsynaptic, or a heterosynaptic element. By analogy, one may distinguish the site at which a work order is issued from the site where the work is done. The work of modifying synaptic strength could be accomplished by the postsynaptic neuron, for instance by altering properties of postsynaptic receptors such as number, turnover, or sensitivity. Alternatively, the postsynaptically issued work order could be carried out presynaptically by altering the amount of transmitter release. Most current evidence in various *Aplysia* ganglia is consistent with a presynaptic site for mechanisms producing synaptic plasticity (Bailey and Chen 1983; Dale et al.



The Site Where Synaptic Strength is Specified and the Site of the Mechanism for Change May be Different

Figure 3.13 Graphic characterization of distinction between site of specification of synaptic strength and the mechanism that varies so as to alter synaptic strength. (Top) Synaptic strength could be specified by the postsynaptic neuron or by the presynaptic neuron; both more global and individual specification are also possible, but are not shown. The pairing data of figure 3.12 are consistent with postsynaptic, rather than presynaptic, specification of synaptic strength in the buccal ganglion network. (Bottom) The mechanism altering synaptic strength, although specified postsynaptically, could be located in either presynaptic or postsynaptic neuron.

1988; Kandel et al. 1987; Mayford et al. 1992; Piomelli et al. 1987; for a review of both presynaptic and postsynaptic contributions, see Pieroni and Byrne 1992). These findings suggest the possibility that postsynaptic specification of synaptic strength involves control of a remote mechanism located in presynaptic terminals (Gardner 1990). This last possibility is explored below in detail, because it appears to require retrograde information transfer across biological synapses that has implications for the applicability of neural network models.

POSTSYNAPTIC NEURONS SPECIFY PRESYNAPTIC QUANTAL RELEASE

The data presented in this section provide a test of the hypothesis that postsynaptically specified differences in synaptic strength reflect differences in presynaptic transmitter release. A corollary of this hypothesis is that synapses formed by different presynaptic neurons converging on

a common postsynaptic cell have similar strengths because the different presynaptic terminals release similar amounts of neurotransmitter: similar numbers of *quanta*.

Assumptions and Methods of Quantal Analysis

Del Castillo and Katz (1954) first observed, at frog neuromuscular junction, that neurotransmitter molecules were released not individually, but instead packaged in multimolecular, similar-sized quanta. Experiments described in this section follow their analysis, as applicable to central nervous system synapses. The basis of the analysis is that successive stimuli release similar but not identical numbers of quanta, and comparison of the mean and variance of sets of synaptic currents can yield indirect measures related to the number of quanta released and the postsynaptic response produced by a single quantum.

Direct quantal analyses, as performed by Del Castillo and Katz (1954), require direct measurements of the postsynaptic response produced by individual quanta. However, at some *Aplysia* synapses, the number of quanta released, the *quantal content*, is large. As a consequence of this and of synaptic convergence, which mixes synaptic noise from several presynaptic inputs, direct measurement of the conductance change produced by single quantum has not been possible. Quantal analyses must therefore be indirect, as above, and this need for indirection increases dependence on particular models. Indirect quantal analyses have been carried out previously in buccal (Baux and Tauc 1987; Chiel et al. 1988; Simonneau et al. 1980) and abdominal (Castellucci and Kandel 1976; Dale et al. 1988) ganglia of *Aplysia*, and have ascribed increases in synaptic efficacy to increased quantal transmitter release. Similar indirect quantal analyses have recently been utilized for the study of hippocampal long-term potentiation (LTP). Objections that have been raised to these analyses in hippocampus on the bases of hypothetical variable terminal activation, receptor saturation, and variability of quantal response are less likely to apply to the *Aplysia* data; for a discussion of these issues, see Gardner (1991, 1993b).

Assurance that transmitter release in *Aplysia* follows particular models would permit deriving values of mechanistic quantal factors such as the probability of release p , the quantal conductance change γ , and the number of release sites n from the data. Instead, imposing fewer constraints on the model for transmitter release yields less information, but still permits tests of the relevant hypotheses. The assumptions underlying this model are detailed in Gardner (1991, 1993b). These can be summarized briefly: (1) transmitter is released quantally, (2) quanta are released independently at a single ending, or at multiple endings that are statistically equivalent to a single ending, (3) statistical parameters underlying quantal release obey stationarity, and (4) the variability in the number of quanta released per action potential is substantially

more than the variability in the number of molecules in each quantum, or in the number of available receptors.

Given conformance to this model, the mean and variance of synaptic current amplitudes can be analyzed to derive quantal parameters descriptive of the synapse. One quantal parameter depends on transmitter release processes localized to the presynaptic terminal; the other parameter is proportional to postsynaptic sensitivity. These factors, obtained for each synapse, can be compared to the strength of that synapse. Moreover, they can be paired by common presynaptic or postsynaptic neuron in order to test the hypothesis that presynaptic transmitter release is partially specified by the postsynaptic neuron.

Each of these factors is derived from the experimentally determined mean M and variance σ^2 of sets of synaptic conductances recorded at a single synapse. Derivations are provided in Gardner (1991, 1993b) and summarized here. The mean conductance is:

$$M = n p \gamma \quad (3.5)$$

with n , p , and γ defined above. Similarly, the variance σ^2 of synaptic conductance is:

$$\sigma^2 = n p \gamma^2 (1 + (CV_\gamma)^2 - p) \quad (3.6)$$

with CV_γ the coefficient of variation of the individual quantal response. These values are then used to derive the first of the two quantal factors of M :

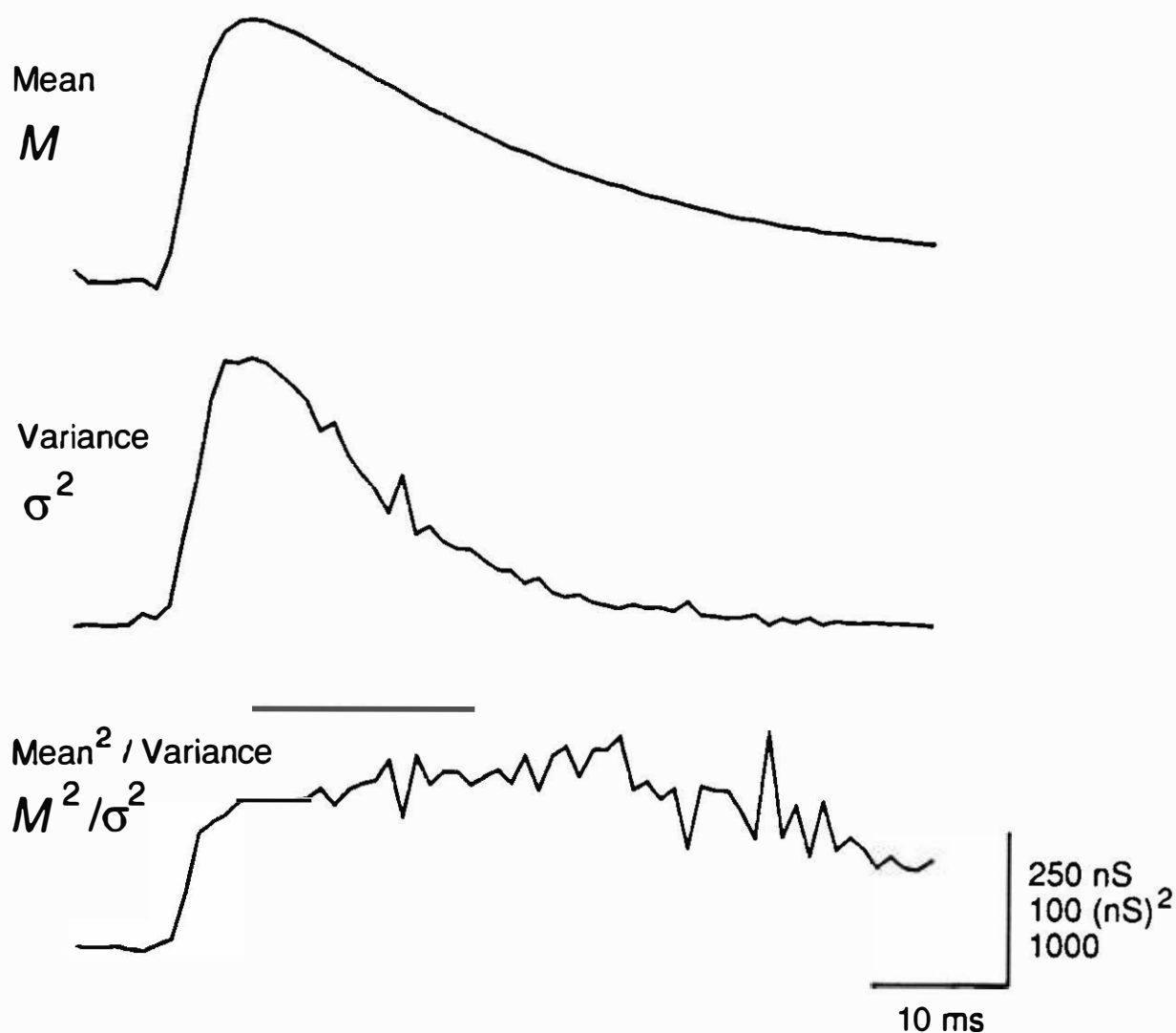
$$\frac{M^2}{\sigma^2} = \frac{n^2 p^2 \gamma^2}{n p \gamma^2 (1 + (CV_\gamma)^2 - p)} = \frac{m}{(1 + (CV_\gamma)^2 - p)} \quad (3.7)$$

where $m = np$ is the quantal content: the number of quanta released per action potential. Dividing M by M^2/σ^2 yields the other quantal factor of M , σ^2/M :

$$\sigma^2/M = \gamma (1 + (CV_\gamma)^2 - p) . \quad (3.8)$$

The factor M^2/σ^2 is essentially presynaptic, dependent on processes located in the presynaptic terminal, and proportional (but not identical) to quantal content m . In contrast, the other factor, σ^2/M , is largely postsynaptic and is proportional to the postsynaptic quantal conductance change γ . The quantal analysis is thus designed to identify presynaptic and postsynaptic factors, permitting each to be correlated with synaptic strength. Although the term $(CV_\gamma)^2$, representing variance of quantal size, can be influenced by postsynaptic microstructure, the term is largely presynaptic (Bekkers et al. 1990). Its presence in the denominator of M^2/σ^2 does not change the identification of this factor with presynaptic rather than postsynaptic processes; possible contributions of $(CV_\gamma)^2$ to changes in M are discussed below.

Figure 3.14 shows derivation of M^2/σ^2 from experimentally recorded mean and variance of sets of PSCs at a single synapse. The top trace



Mean, Variance, and M^2/σ^2 of a Set of Postsynaptic Currents

Figure 3.14 The result of ensemble averaging 31 successively recorded postsynaptic currents at a single synapse. The process yields an index, M^2/σ^2 , that reflects presynaptic quantal factors and reveals the time course of the variance to be consistent with quantal models. (Top) The mean postsynaptic current shows the time course of the underlying conductance M . (Middle) The ensemble variance σ^2 is near zero preceding the PSC, with approximate time course of the square of the synaptic conductance. (Bottom) M^2/σ^2 calculated from top and bottom traces is equal to $m/(1 + (CV_y)^2 - p)$, and thus reflects essentially presynaptic factors (see text). The horizontal bar shows the 16-ms interval following the PSC peak during which the average value of M^2/σ^2 was computed. For each trace, only the initial 64 ms is shown. (Reprinted from Gardner 1991, with permission of the American Physiological Society.)

shows the ensemble average of 31 successively recorded PSCs. The peak of this average conductance is M , used as index of the static synaptic strength characterizing the synapse. The variance σ^2 of the set of PSCs is shown as the middle trace. The quantal parameter M^2/σ^2 , derived from M and σ^2 , is shown as the bottom trace. In order to reduce fluctuations of individual time points, the value of M^2/σ^2 used for subsequent analyses and comparisons was obtained by averaging this trace for 16 ms following the PSC peak.

There exist two potential sources of artifact in this analysis. Calculating quantal parameters using large numbers of synaptic currents produces more reliable values for statistical parameters only if the data set is truly stationary. Because biological samples rarely display long-term stationarity, and because artifacts of correlated drift in synaptic record-

ings would be likely to produce false correlations supporting the presynaptic hypothesis, sample sizes were intentionally kept small (Gardner 1991, 1993b; Malinow 1991). However, analyses of longer runs produced essentially similar correlations and conclusions.

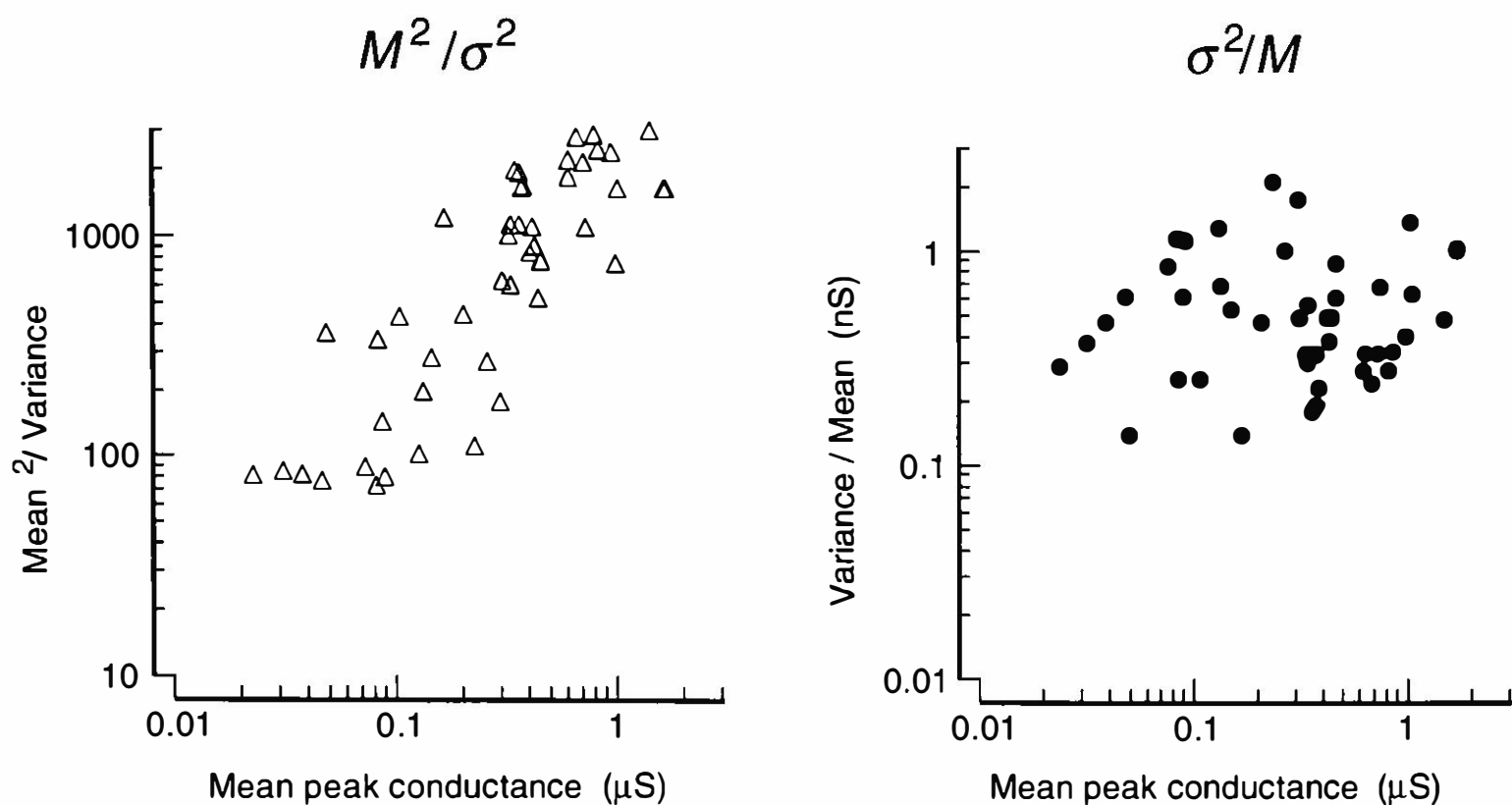
The second potential artifact could contribute to the observed variance in M . For smaller values of M , variations due to individual quanta become increasingly detectable, permitting the comparative evaluation of underlying events that forms the basis of the quantal analysis. However, membrane noise might serve as an additional source contributing to the increasing variation of smaller PSCs. Ascribing all such variance to quantal sources would distort the analysis and produce artificially low values of M^2/σ^2 and abnormally high values of σ^2/M , again falsely supporting the presynaptic hypothesis. By determining the PSC-to-PSC variance for much of the time course of the recorded PSC, as shown in figure 3.14, rather than just for the single value of peak conductance, it is possible to exclude this artifact as a significant contributor to the derived quantal parameters. By most models, membrane noise should contribute time-independent variance. Instead, because the observed variance follows the approximate time course of the square of the PSC, it is revealed to be of synaptic origin (Gardner 1991, 1993b).

Synaptic Strength Reflects Presynaptic Transmitter Release

This analysis yields values of M and of its two quantal factors M^2/σ^2 and σ^2/M for each synapse. Since the two quantal factors represent largely presynaptic and partially postsynaptic contributions, relating M to its factors can provide information as to whether presynaptic or postsynaptic processes are varied in order to increase or decrease synaptic strength. Figure 3.15 compares M , as a measure of synaptic strength, to each of its two factors. In each graph, the horizontal axis plots M for each of 46 synapses. In the left panel, M^2/σ^2 is well correlated with M , consistent with the hypothesis that synaptic strength reflects presynaptic factors. The right panel, in contrast, shows no correlation between σ^2/M and M , showing that synapses of different efficacy cannot be the result of systematically varying this partially postsynaptic factor. The dependence of synaptic strength on the essentially presynaptic factor, and lack of dependence on the factor including postsynaptic sensitivity, is confirmed by analysis of variance. The relationship between M^2/σ^2 and M is significant ($P = .0001$); the relation between σ^2/M and M is not ($P = .64$) (Gardner 1991).

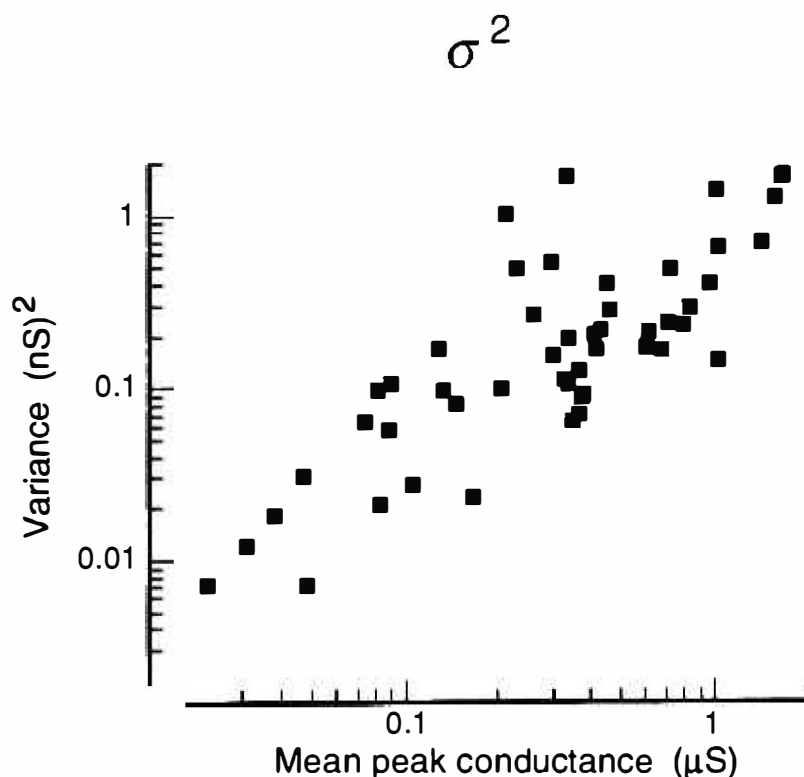
In contrast to the more limited range of potentiation observed dynamically at hippocampal LTP, the static values of both synaptic strength M and approximate quantal content M^2/σ^2 in the *Aplysia* population varied widely, with M^2/σ^2 ranging from 70 to nearly 3000. In contrast, σ^2/M , which was independent of M , showed less variation.

Additional insight into both interpretation of M^2/σ^2 and identification



Peak Conductance Reflects a Presynaptic Mechanism

Figure 3.15 Synapse-to-synapse differences in mean peak conductance M reflect differences in the presynaptic factor M^2/σ^2 . For each of 46 synapses, values of M^2/σ^2 (Δ) and of σ^2/M ($\bullet$) are plotted against the corresponding values of M . (Left) M^2/σ^2 is well correlated with M ($r = .85$, $P = .0001$ for the regression line of near-unity slope on this log-log plot). (Right) The factor σ^2/M , which reflects postsynaptic sensitivity as well as some presynaptic contributions, is uncorrelated with M ($r = .07$, $P = .64$). (Reprinted from Gardner 1991, with permission of the American Physiological Society.)



Variance Does Not Decrease at Higher Means

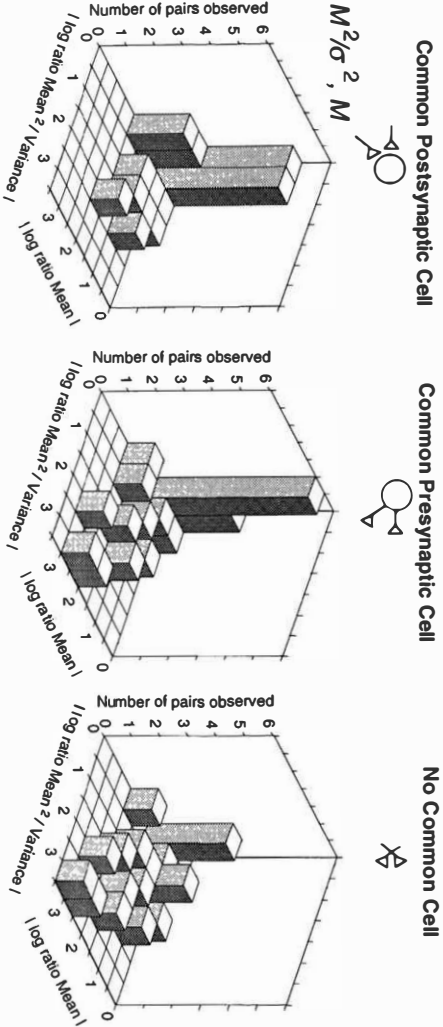
Figure 3.16 Variance and mean of postsynaptic conductance are correlated, consistent with low values of quantal probability p . For each of the synapses shown in figure 3.15, values of variance σ^2 are plotted against the corresponding values of M ($r = .81$, $P = .0001$ for a regression line of slope 1.02 on this log-log plot).

of sources of quantal variance is provided by plotting σ^2 as a function of M for each of the synapses in the population (figure 3.16). The pre-synaptic factor M^2/σ^2 approximates quantal content m as p approaches zero, and for small values of $(CV_\gamma)^2$. If increases in p or decreases in $(CV_\gamma)^2$ contributed significantly to larger M , figure 3.16 would show pronounced downward curvature toward the right. Instead, the approximately linear relation suggests that p remains low and $(CV_\gamma)^2$ constant. M^2/σ^2 depends inversely on $(CV_\gamma)^2$ as well as directly on quantal content m , so for the observed increase in M^2/σ^2 with increasing M to reflect a decrease in $(CV_\gamma)^2$ rather than an increase in m , it would be necessary to postulate an implausible mechanism in which more precise packaging of transmitter was well correlated with greater postsynaptic responsivity over more than an order of magnitude.

In other systems, quantal variance has been ascribed, not to release of varying numbers of quanta of neurotransmitter, but to varying numbers of all-or-none presynaptic terminals, or to postsynaptic quantization of a small number of receptors (Edwards et al. 1990; Redman 1990). The lack of downward curvature of figure 3.16 at larger M may make some of these alternate interpretations less likely; other data reinforcing the relation between quantized M^2/σ^2 and transmitter release are reviewed in Gardner (1991, 1993b).

This section has presented data showing that synaptic strength of synapses reflects values of the essentially presynaptic quantal factor M^2/σ^2 . The previous section compared pairs of synapses to show synaptic strengths to be similar at convergent synapses, suggesting postsynaptic specification of synaptic strength. To complete this analysis, I present paired comparisons of both M and M^2/σ^2 to show that the similar strength of synapses made by convergent terminals can be ascribed to their similar values of presynaptic M^2/σ^2 . The dual-parameter histograms of figure 3.17 compare similarities of M and M^2/σ^2 at paired synapses. For each parameter, similarities are shown by calculating the absolute value of the natural logarithm of the ratio of the two paired values at each synaptic pair. This procedure yields results closer to zero for more similar pairs. Histograms grouping values of these absolute log ratios for both M and M^2/σ^2 are then plotted for each of the three pairing schemes of figure 3.11. Greater conjoint similarity of both parameters is represented by bars closer to the origin, shown at the far center of each plot. The common postsynaptic cell case shown in the left graph of figure 3.17 shows that synapses converging on a common postsynaptic neuron have similar strength because they have similar values of the essentially presynaptic factor M^2/σ^2 . All three plots show values clustering along the horizontal diagonal, consistent with the correlation of M and M^2/σ^2 shown for unpaired synapses in figure 3.15.

Not shown here are pairings for the largely postsynaptic parameter σ^2/M . All pairings for this parameter are close and indistinguishable,



suggesting that systematic variation of σ^2/M does not contribute to similarities or differences in synaptic strength at paired synapses.

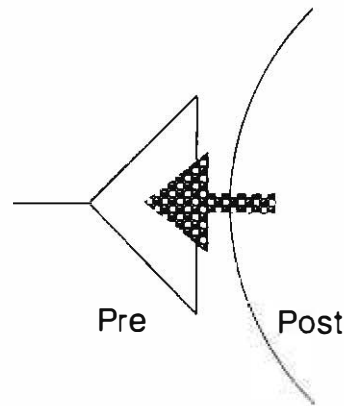
The data thus ascribe differences in synaptic amplitude to varying numbers of quanta released by presynaptic terminals. Similar synaptic strengths observed at convergent synapses of different presynaptic cells on common postsynaptic neurons result from the release of similar numbers of quanta. The findings show both synaptic strength and presynaptic transmitter release to be partially specified by postsynaptic neurons.

APLYSIA NEUROBIOLOGICAL MECHANISMS CONSISTENT WITH RETROSYNAPTIC INFORMATION TRANSFER

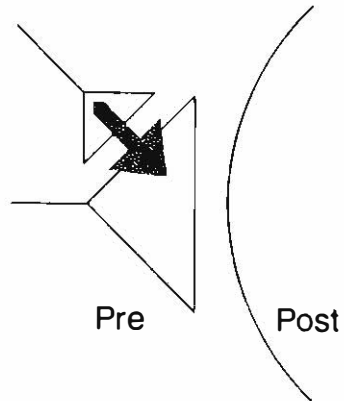
Individual buccal ganglia interneurons release different amounts of neurotransmitter at divergent terminals, with these amounts partially determined by the postsynaptic target cell. These quantitative differences contrast to the qualitatively similar postsynaptic currents received by these postsynaptic cells, most of which are motor neurons with similar function. Moreover, synaptic strength and quantal release vary together at the same identified synapses from animal to animal. Again, this quantitative variation is unlike other properties of identified neurons, making it unlikely that different synaptic morphology is determined exclusively by genetically programmed development. As noted in part I, these findings suggest that quantitative differences in synaptic efficacy do not result from fixed genetic or developmental characteristics of particular postsynaptic neurons, but may instead arise postdevelopmentally, from plastic modulation of synaptic anatomy or physiology.

According to this interpretation, the differences seen in M represent the result of long-term, possibly adaptive, modification of synaptic strength, specified by postsynaptic targets but expressed by modulation of transmitter release. In order to implement such a process, information must be transferred in a retrograde fashion, from postsynaptic to presynaptic neuron. Figure 3.18 presents three classes of mechanisms consistent with apparent retrograde information flow, each of which could plausibly mediate long-term postdevelopmental specification of presynaptic release by postsynaptic neurons. These are: (A) retrograde synaptic transmission by a reverse transmitter, (B) modulatory heterosynaptic transmission, and (C) the postdevelopmental alteration of presynaptic terminal morphology by postsynaptic targets. These mechanisms have been selected in part because there exists partial evidence in *Aplysia* for processes that could underlie parts of each, and evidence as well for more than one candidate neurochemical mediator.

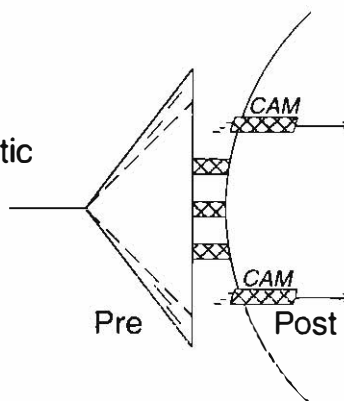
A. Retrosynaptic transmission



B. Heterosynaptic modulation



C. Control of presynaptic terminal expansion



Hypothetical Mechanisms for Retrograde Information Transfer

Figure 3.18 Three mechanisms for retrograde information transfer consistent with postsynaptic specification of presynaptic quantal release, shown schematically. (A) Retrosynaptic transmission of a chemical mediator which is released by the postsynaptic neuron and which modulates release of the conventional neurotransmitter by presynaptic terminals. (B) Heterosynaptic modulation of presynaptic terminals by an additional neuron receiving input fed back from the postsynaptic cell. (C) Control of presynaptic terminal expansion, and therefore efficacy, indirectly by up- or down-regulation of postsynaptic cell-adhesion molecules. See text for details.

Retrosynaptic Transmission

According to the first class of mechanisms, differences in synaptic strength result from the action of retrosynaptic modulators, perhaps diffusible substances released selectively by postsynaptic neurons and modulating subsequent transmitter release from each of the incident presynaptic terminals within their immediate range.

Putative retrosynaptic transmitters in this system include arachidonic acid and its metabolites, and the neurotransmitter acetylcholine itself. In *Aplysia*, arachidonic acid and its metabolites have been investigated both as transcellular messengers and as intracellular second messengers (Feinmark et al. 1989; Piomelli et al. 1987). Selective release of arachidonic acid or its metabolites such as 12-HPETE is thus a plausible mechanism for postsynaptic control of terminals of B4 and B5. Arachidonic acid has also been proposed as a retrograde transmitter in hippocampus (Williams et al. 1989).

A different form of retrosynaptic information flow in these ganglia, in the self-inhibitory synapses that both B4 and B5 make upon themselves, suggests an additional transsynaptic candidate mechanism. Both nicotinic and muscarinic cholinergic autoceptors are found on B4 and B5, modulating presynaptic release (Fossier et al. 1988; Gardner 1977a). Indeed, Fossier et al. (1988) reported that potentiation of transmitter release following treatment with muscarinic agents persisted well past washout, consistent with activation of a long-term process. The retrograde synaptic control, specific to presynaptic terminals on individual postsynaptic neurons, might therefore be mediated by acetylcholine or even by choline produced by local hydrolysis, which demonstrates limited neuroactivity at these synapses (Gardner et al. 1984). The requirement for postsynaptic specificity might be satisfied by differing amounts of cholinesterase localized at individual postsynaptic neurons, or by locally varying barriers to diffusion.

Heterosynaptic Modulation

The second class of putative mechanisms reflects a heterosynaptic rather than a direct postsynaptic influence. According to this class, presynaptic differences in transmitter release are specified by varying amounts of modulation by such agents as serotonin or histamine released differentially on synaptic terminals by a heterosynaptic connection (Chiel et al. 1988; Glanzman et al. 1989, 1990; Goldstein et al. 1984). Such modulators can differentially and independently facilitate divergent terminals of a single *Aplysia* neuron (Clark and Kandel 1984.) Many additional substances have been identified as modulatory transmitters in *Aplysia*, including agents that decrease and others that increase synaptic strength. In buccal ganglia alone, the small peptides SCP_B and FMRFamide excite B4 and B5 and modulate their followers so as to produce modified activity (Sossin et al. 1987); histamine depresses synaptic transmission

and reduces M^2/σ^2 , consistent with a presynaptic mechanism (Baux et al. 1990; Chiel et al. 1988). Analyses and reviews of multiple presynaptic effects of various neuromodulators including buccalin and myomodulin in *Aplysia* ganglia have been presented by Rosen et al. (1989), Pieroni and Byrne (1992), and Weiss et al. (1992).

The mechanisms of action of heterosynaptic neuromodulators have also been studied in *Aplysia*. In abdominal ganglia, transmitter release and synaptic plasticity have been shown to depend on two processes, one affecting Ca^{2+} transients and one regulating steady-state Ca^{2+} . The modulatory neurotransmitters serotonin and FMRFamide modify synaptic strength in opposite directions by modulation of transient I_{Ca} , whereas presynaptic voltage manipulations can modulate steady-state Ca^{2+} (Blumenfeld et al. 1990; Connor et al. 1986; Shapiro et al. 1980). These different mechanisms alter different Ca^{2+} channels: dihydropyridine (DHP)-insensitive for classical plasticity and DHP-sensitive for tonic V_m changes (Edmonds et al. 1990). In addition, various other voltage-dependent channels are modulated by serotonin and by peptides, in some cases via known second messenger pathways. Heterosynaptic mechanisms are related to classes of learning rules in chapter 4.

Morphological Retrosynaptic Modulation

A third possible mechanism for retrograde information transmission postulates postsynaptic control of presynaptic morphology. In *Aplysia*, several studies correlate synaptic strength with the number of presynaptic release sites and the extent of branching. For example, postdevelopmental habituation or sensitization produces plastic modulation of the number of presynaptic active zones in the abdominal ganglion (Bailey and Chen 1983) and the number of presynaptic varicosities per sensory neuron in the stimulated pathway (Bailey and Chen 1988). In culture, serotonin facilitates transmission by means of an increase in apparent quantal content, as determined both by the method of failures and by comparison of EPSPs and miniature EPSPs, possibly due to an increase in the number of presynaptic release sites (Dale et al. 1988). Presynaptic plasticity depends in part on postsynaptic influence. During development in vitro, the presence of appropriate *Aplysia* postsynaptic neurons modulates presynaptic morphology so as to affect synaptic efficacy (Glanzman et al. 1989). Similarly, serotonin-dependent functional and structural facilitation were found to depend on interaction of the presynaptic terminals with appropriate postsynaptic targets (Glanzman et al. 1990).

Plausible mechanisms also exist by which presynaptic morphology, and consequently synaptic strength, might be modified by the postsynaptic target. In these initial studies, plasticity was induced by applied serotonin, analogous to a heterosynaptic modulator. Serotonin application produced internalization and down-regulation of presynaptic cell-

adhesion molecules (CAM), producing synaptic plasticity by removing inhibition on growth of presynaptic structures (Bailey et al. 1992; Mayford et al. 1992). The internalization observed appears to consist of membrane components from both neurons, suggesting the involvement of cell-cell signaling in this process (Bailey et al. 1992). Peter et al. (1992) found that the neuromodulator FMRFamide down-regulated postsynaptic *Aplysia* CAM, affecting neurite outgrowth. Postsynaptic neuron-dependent neuromodulation of the buccal synapses, therefore, might possibly be mediated through an alteration in the amount of postsynaptic *Aplysia* CAM, in turn regulating changes in presynaptic terminal growth. For further review of the role of CAM in synaptic modification and plasticity, see Schubert (1991); polysialic acid has also been implicated (Rutishauser and Landmesser 1991).

The dependence of each of these hypothesized presynaptic mechanisms on the postsynaptic neuron should be noted. Activity-dependent long-term changes in protein kinase, known to be responsible for mechanisms such as long-term plasticity at several sites in *Aplysia* (Kandel et al. 1987), would be expected to influence all presynaptic terminals of a given neuron. However, the independent synaptic strengths of individual terminals would not be expected from a homosynaptic activity-dependent mechanism; divergent terminals share a common history, yet can mediate PSCs of very different mean amplitude.

Inclusion of the particular hypothetical mechanisms discussed above has been based on at least partial evidence for each in *Aplysia*. Other mechanisms can be postulated, including ephaptic interactions between presynaptic and postsynaptic specializations, and selective modulation of transmission of action potentials through presynaptic branches, but evidence for these is lacking.

STATIC AND DYNAMIC RETROSYNAPTIC PLASTICITY IN NEUROBIOLOGY

The indirect evidence favoring a retrosynaptic or heterosynaptic process in determination of static synaptic strength in *Aplysia* buccal ganglia would be strengthened by direct results supporting such a mechanism in other neuronal networks. Although there now exist several sites for which partial evidence is available, the phenomenon of long-term potentiation (LTP) in mammalian hippocampus provides the clearest evidence of any preparation for the existence of a retrosynaptic messenger mediating synaptic plasticity.

Retrosynaptic Mechanisms of Hippocampal LTP

Hippocampal LTP is a neuronal analog of memory in which afferent input is potentiated for several hours or more. Many groups are actively engaged in elucidating the underlying mechanisms, which probably

differ at distinct hippocampal sites. There is general agreement that induction of LTP in the CA1 region requires activation of a particular type of postsynaptic glutamate receptor named for its selective activation by *N*-methyl-D-aspartate (NMDA), producing a localized increase in postsynaptic intracellular Ca^{2+} . Note that chapters 2 and 4 discuss additional aspects of the relation of NMDA-mediated synapses to neural network models. For reviews of LTP, see Baudry and Davis (1991) and Bliss and Collingridge (1993).

In spite of this clear requirement for postsynaptic intracellular changes to specify induction, longstanding evidence points to increased presynaptic transmitter release as the mechanism for maintained potentiation of synaptic strength (Dolphin et al. 1982; Skrede and Malthé-Sørensen 1981). Coupling increased glutamate and aspartate release to the postsynaptic dependence of associative induction, Bliss et al. (1986) postulated the need for a retrosynaptic signal from postsynaptic cells to act on recently activated presynaptic terminals.

Likely candidate molecules for the retrosynaptic messenger were quickly proposed. Stimulation by glutamate of cerebellar neurons containing NMDA induced Ca^{2+} -dependent release of nitric oxide (NO), leading Garthwaite, Charles, and Williams (1988) to propose NO as the hippocampal transsynaptic factor. The additional utilization of arachidonic acid or its lipoxygenase metabolites was noted by Williams et al. (1989), who suggested that NO might mediate the early retrosynaptic message, while arachidonic acid served as retrograde messenger for later components of LTP. Theoretical legitimization of the prior experimental findings was provided by Gally et al. (1990), who modeled several developmental and plastic consequences of this plasticity-inducing diffusible substance.

Given postsynaptic NMDA-dependent induction, evidence for potentiation by enhanced transmitter release rather than increased postsynaptic receptor sensitivity emphasizes the requirement for a retrosynaptic modulator. Independently, Malinow and Tsien (1990) and Bekkers and Stevens (1990b) employed quantal analysis to identify the majority of LTP expression as due to enhanced presynaptic release. In spite of additional postsynaptic contributions (Manabe et al. 1992) and alternate hypotheses for the interpretation of quantal data (Larkman et al. 1991; but see Larkman et al. 1992), this work greatly strengthened the evidence for a presynaptic target site for a transsynaptic messenger.

Further evidence for the role of NO in LTP was provided by O'Dell et al. (1991) and Schuman and Madison (1991), who reported independently that intracellular injection of inhibitors of NO synthetase into postsynaptic cells prevented LTP, whereas extracellular application of NO-binding hemoglobin reduced LTP. O'Dell et al. (1991) additionally found that delivery of nitric oxide enhanced transmitter release from cultured hippocampal cells. However, carbon monoxide has also been

proposed as a retrosynaptic messenger and precise identification of the messenger and verification of its synthesis is lacking at present.

Retrosynaptic Mechanisms at Other Sites

Retrosynaptic information transfer has been identified or postulated at sites other than hippocampus or *Aplysia*. Postsynaptic trophic modulation of presynaptic morphology and function has been proposed to mediate competition between convergent inputs, stabilization or rejection of innervation, and regulation of neuronal function in response to changes in the size or form of the organism (for reviews, see Changeux and Danchin 1976; Purves et al. 1988).

Retrosynaptic modulation has been described in other preparations, in some of which plastic and trophic effects are combined. In *Xenopus* nerve-muscle co-cultures, postsynaptic increases in Ca^{2+} within myocytes produced a decrease in ACh release from nerve terminals, unless the terminals were protected by Hebbian synchronous co-activation. This depression is conjectured to involve a transient and transsynaptic retrograde factor (Dan and Poo 1992). Collins et al. (1984) compared plasticity and presynaptically-determined EPSP size to suggest that motor neurons indirectly control the probability of transmitter release from Ia afferents innervating them.

Retrosynaptic modulation different in mechanism and function, resembling in part the scheme described in figure 3.18B, has been described in crab stomatogastric ganglion, where several target neurons feed back to modulate specific presynaptic terminals of innervating fibers. This dynamic interaction is conjectured to selectively alter synaptic efficacy by differentially affecting release of peptide and conventional neurotransmitters (Nusbaum et al. 1992).

RETROSYNAPTIC MECHANISMS FOR NETWORK LEARNING RULES

The analyses presented in this chapter, unlike the hippocampal LTP protocols reviewed, are static, in that differing synaptic strengths are examined, rather than altered. We view what is presumed to be the result of dynamic processes that have contributed to the setting of synaptic strength. As a result, we have no knowledge of the learning rule involved, or if the rule is modified at different terminals of the same presynaptic neuron, or even if the changes in M observed result from application of a single learning rule at different synapses.

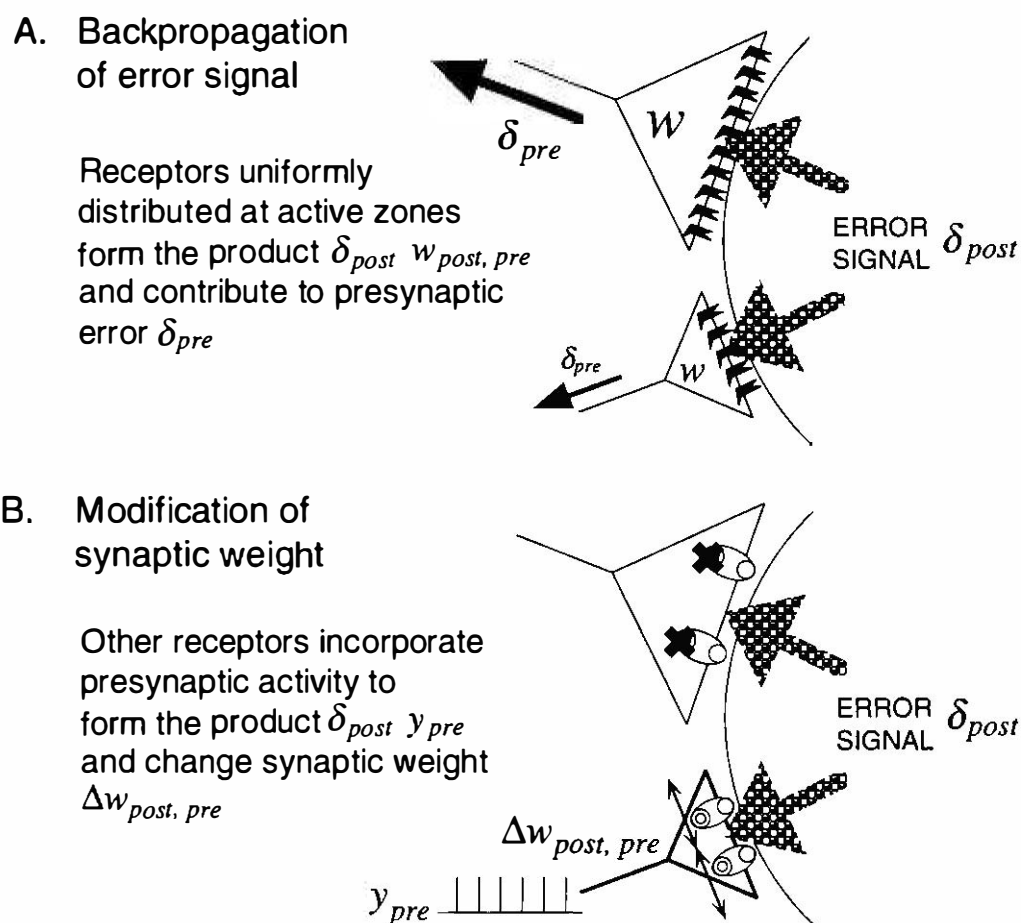
With this caution, it is possible to close this sample attempt to synthesize neurobiology and neural network models by proposing cellular and synaptic neurobiological mechanisms that could be utilized to implement the requirement for retrograde information transmission imposed by backpropagation.

Neurobiological Requirements for Backpropagation

As reviewed at the start of part II, backpropagation requires the retrosynaptic transmission of an error signal specified by the postsynaptic neuron. Presynaptic units utilize this signal in two different ways, and in each of these the error is scaled independently. In one, the error signal itself, scaled by the synaptic strength, is backpropagated to contribute to the error signal characterizing the presynaptic cell. In the other, the error scaled by presynaptic activity is used to adjust local synaptic weights. A consistent interpretation of these processes is that in each case, the postsynaptic neuron transmits a global error signal characteristic of the neuron, rather than individual synapses. Individual presynaptic terminals receive this message in two distinct ways, one local to the synapse and the other, dependent on activity, presumably characterizing the presynaptic neuron.

Examples of neural mechanisms outlined above drawn from both *Aplysia* and hippocampus could be combined to mediate these actions. Diagrams of these hypothesized syntheses are shown in figure 3.19. Both direct evidence in hippocampus and indirect inferences from *Aplysia* are consistent with postsynaptic release of neuroactive substances. In hippocampus, release of the retrosynaptic transmitter is local to the postsynaptic specialization, whereas evidence from paired presynaptic endings in *Aplysia* ganglia suggest a more general postsynaptic release. Retrosynaptic transmission requires not only postsynaptic release but also presynaptic receipt. Again, evidence from hippocampus points not only to a requirement for postsynaptic release of the putative messenger NO, but also to the ability of diffusible NO to induce changes in presynaptic release (O'Dell et al. 1991). *Aplysia* data support changes to static synaptic strength that are performed by presynaptic mechanisms under the influence of postsynaptic specification. According to the backpropagation model, the signal to be transmitted retrosynaptically is specific to the postsynaptic cell, and varying action at convergent synapses follows distinctions between presynaptic neurons or between individual presynaptic terminals, rather than between regions of the same postsynaptic cell. As noted above, patterns of *Aplysia* synaptic strength are consistent with this picture. Some data from hippocampus similarly suggest an activity-dependent mechanism, and these are reviewed below in the discussion of Hebbian mechanisms.

The two distinct roles of the postsynaptic error signal, diagrammed in figure 3.19, could be mediated by two classes of presynaptic receptors, one to transform the error and transmit it back to lower-order cells, and one to modify the strengths of synapses at that level (Gardner 1993a). To implement the first of these functions, backpropagation of the error signal, presynaptic terminals would need receptors in proportion to the strength of each synapse. If increases in synaptic strength reflected increases in presynaptic terminal area, then larger presynaptic



Hypothetical Retrosynaptic Mechanisms for Backpropagation

Figure 3.19 Hypothetical neurobiological mechanisms for backpropagating both error signal and weight change. In each case, the postsynaptic neuron is postulated to release a retrosynaptic chemical mediator proportional to its error signal δ_{post} . (A) The error signal to be integrated by the presynaptic neuron is captured by a larger number of uniformly spaced receptors on larger presynaptic terminals, satisfying the requirement that the backpropagated error δ_{pre} be weighted by the forward synaptic strength $w_{post,pre}$. (B) Local synaptic weight change $\Delta w_{post,pre}$ is mediated by a different receptor mechanism, which responds to δ_{post} only when the presynaptic terminal is currently or recently depolarized ($y_{pre} > 0$), consistent with the activity-dependence requirement for weight change. Compare to figures 3.7 and 3.8.

terminals could have greater numbers of uniform-density receptors specialized to mediate this backpropagation of the postsynaptic signal. The retrosynaptic signal acting on these receptors would be scaled by synaptic strength, forming the appropriate product to contribute to error propagation.

The weight-change function of the postsynaptic error signal, in contrast, could be mediated by a distinct class of presynaptic receptors whose response depended on not only the retrosynaptic transmitter, but also on an index of presynaptic activity, such as local Ca^{2+} concentration or depolarization of the terminal. This latter class of receptors would thus be specialized to form the appropriate product to alter synaptic weights. One such mechanism, activity-dependent neuromodulation, is discussed in chapter 4. Unlike that of the first class of receptors, the response of these receptors would itself be independent of presynaptic terminal area or its analog synaptic strength.

The utilization of two distinct receptors to mediate different conse-

quences of the same postsynaptic error signal may be summarized by relations derived from equations 3.3 and 3.4. Each postsynaptic cell backpropagates its error δ_{post} in proportion to synaptic weight $w_{post,pre}$ to form the error for the presynaptic cell:

$$\delta_{pre} \propto \sum_{post} \delta_{post} w_{post,pre} . \quad (3.9)$$

The same δ_{post} scaled by input activity y_{pre} alters $w_{post,pre}$:

$$\Delta w_{post,pre} \propto \delta_{post} y_{pre} . \quad (3.10)$$

In their alternative view of plausible neural mechanisms for learning models, Zipser and Rumelhart (1990) have already noted the computational advantages of separating responses to a common neuromediator, drawing analogies to the use of multiple postsynaptic receptors to glutamate. Dual receptors to neuromodulators affecting transmitter release have been identified and characterized. B4 and B5 each contain both nicotinic and muscarinic autoceptors to ACh, each of which modulates transmitter release differently (Fossier et al. 1988); these cells also respond differently to FMRFamide and 5-HT (Baux et al. 1990). Elsewhere in *Aplysia*, dual cholinergic receptors affecting transmitter release were identified as early as 1975 (Woodson et al. 1975). Additionally, the presynaptic modulatory agents FMRFamide and 5-HT mediate two distinct and separable effects, spike broadening and transmitter mobilization; in addition, each has both short-term and long-term effects (Baxter and Byrne 1990; Pieroni and Byrne 1992). SCP_B , FMRFamide and 5-HT each mediate their effects through multiple receptors and multiple second messengers or processes (Taussig et al. 1990). These distinctions can be functional as well as pharmacological. The facilitatory neuromodulators, SCP_B and 5-HT, have differential effects on depressed and normal synapses, due to differing requirements for cAMP (Schacher et al. 1990). Neuromodulatory substances can *increase* excitability while *decreasing* transmitter release, and therefore synaptic strength, of the same cell (Whim and Lloyd 1992).

The mechanisms hypothesized in figure 3.19 are specialized to transmit retrosynaptic information across synapses. Backpropagation of error signals requires in addition the conveyance of scaled error signals from presynaptic terminals back to postsynaptic specializations of the same neuron. Electrical signals (in small neurons), Ca^{2+} -dependent processes or axonal transport, among others, could be invoked as possible mechanisms for this partially non-neuromorphic requirement for retrograde intraneuronal transmission. However, *Aplysia* data may provide a simpler plausible biological substrate. These cells are often unipolar, and even when bipolar, the soma does not separate input and output regions. Connections are made in a neuropil surrounded by neuronal cell bodies. In these neurons, regions of presynaptic and postsynaptic specializations are not distinct, and both are electrically near to the

soma. Consequently, even local information arriving retrosynaptically at presynaptic terminals could influence postsynaptic specializations of the same neuron. Similar possibilities are provided by local circuit neurons in both vertebrate and invertebrate nervous systems.

It must be emphasized that although plausible biological mechanisms exist for each hypothesized step in this procedure, I do not in any way mean to suggest that backpropagation is actually employed by any nervous system, merely that it is not possible to reject classes of models or algorithms as nonbiologically based solely on a dependence upon retrograde information transmission.

Neurobiological Requirements for Hebbian Learning Rules

Although computationally powerful, backpropagation is not the only network model learning rule that imposes a requirement for retrograde information transmission. Any Hebbian synapse imposes a requirement for simultaneous presentation of presynaptic and postsynaptic signals at a single locus. If the Hebbian locus is located postsynaptically, the postsynaptic neuron can simultaneously integrate its own activity and that of the presynaptic neuron using information supplied via the conventional neurotransmitter. However, the majority of mechanisms for altering synaptic efficacy appear to be located presynaptically. Hebbian mechanisms at presynaptic loci require retrograde information to be provided by the postsynaptic mechanism, similar to the need imposed by backpropagation. Biological mechanisms mediating Hebbian modification at a presynaptic locus are formally not too different from those that could be invoked to account for backpropagation. Only two distinctions need be made. The retrograde signal codes *activity* for Hebbian learning rules, and *error* for backpropagation, and only backpropagation imposes a requirement for two classes of receptors for the retrosynaptic messenger. The retrosynaptic or heterosynaptic mechanisms I have discussed could mediate the transfer of information from postsynaptic to presynaptic element needed to implement a Hebbian mechanism with expression in the presynaptic terminal.

However, if the differences in M seen result from a Hebbian mechanism, it would have to be a modified mechanism in which postsynaptic neuron activity was in some way weighted more heavily than presynaptic. If presynaptic and postsynaptic activity were equally weighted, or if there existed an all-or-none requirement for near-simultaneous activity in presynaptic and postsynaptic neurons, then both presynaptic and postsynaptic pairings should produce identical similarity. This analysis assumes that divergent presynaptic terminals of the same neuron have shared a common history as well as currently demonstrating identical patterns of activity. In addition, the similar inputs and electrotonic coupling of B4 and B5 introduce sufficient commonality of activity suggesting that if the differences in M seen result from an activity-

dependent mechanism, this must be sensitive to differences between the two cells, or else that their firing patterns were more distinct at a previous time when the plastic changes were implemented.

Neurobiologically Consistent Learning Rules

A classification of learning rules for both excitatory and inhibitory synapses, based on differing contributions of presynaptic and postsynaptic elements, has been proffered by Levy and Desmond (1985). An additional analysis of learning rules from a biological perspective is provided by Baxter and Byrne (chapter 4). The buccal ganglia data presented above would appear consistent with three of the four inhibitory learning rules (1a, 1b, and 2) suggested by Levy and Desmond (1985). Whatever the learning rule or rules, it is impossible on the basis of the data presented to know whether the synaptic strength is increased in response to activity or to inactivity of individual postsynaptic neurons, or of both cells forming the synapse. This distinction between associative and nonassociative rules is extensively analyzed by Baxter and Byrne (chapter 4). In particular, the results are consistent with a postsynaptic growth learning rule, in which activity in individual postsynaptic neurons regulates receptivity to innervation. This learning rule is closely analogous to the mechanism shown in figure 3.18C, in which individual postsynaptic neurons selectively modulate synapse-specific molecules in such a way as to alter the number of active zones in the incident presynaptic terminal.

Similar mechanisms can mediate very different learning rules. For example, one class of learning rule that would not be expected to yield the observed patterns of synaptic strength, in which convergent terminals on a single postsynaptic follower show greater similarity than unpaired inputs, is postsynaptic competition. For this, a reciprocal relation between synaptic strengths of convergent inputs would be a more likely static outcome. However, an activity-dependent Hebbian competition between presynaptic neurites at neuromuscular junction (Lo and Poo 1991) is likely to be mediated by the transsynaptic retrograde mechanism proposed by Dan and Poo (1992).

Another separation between learning rules and mechanisms is provided by consideration of neuronal specificity of long-term potentiation. In spite of convincing association of hippocampal LTP with both Hebbian learning rules in particular and connectionist models in general (Brown et al. 1988), it is not clear whether the apparently retrosynaptically mediated potentiation displays true Hebbian specificity. Kelso et al. (1986) reported Hebbian specificity for stimulated versus unstimulated presynaptic pathways converging on common postsynaptic CA1 pyramidal cells. Recently, Small et al. (1992) found that the putative retrosynaptic messenger NO produced long-term enhancement of only active inputs to CA1, consistent with activity-dependent presynaptic specificity.

Bonhoeffer et al. (1989) found a lack of postsynaptic specificity, with LTP at synapses made by divergent terminals of stimulated presynaptic fibers on unstimulated postsynaptic CA1 neurons. Because of the similar firing patterns of *Aplysia* neurons B4 and B5, the postsynaptic specification of convergent presynaptic inputs found in buccal ganglia and reported in this chapter would seem consistent with the actions of a retrosynaptic diffusible substance on either active terminals alone, or on all incident presynaptic terminals.

ACKNOWLEDGMENT

For support of these investigations, I thank the National Institute of Neurological Diseases and Stroke, National Institutes of Health, for grant NS11555, the Office of Naval Research for contract N00014-86-K-0491 and grant N00014-90-J-1460, and the Air Force Office of Scientific Research for University Research Instrumentation Program grant AFOSR-87-0017.

4

Learning Rules from Neurobiology

Douglas A. Baxter and John H. Byrne

Among the more interesting properties of biological neural networks is their ability to be modified by experience. Many artificial neural networks seek to emulate this ability by including mathematical expressions that allow for experience-dependent modifications of some aspect of the network. Thus, in order both to understand biological neural networks and to simulate their properties, it is necessary to understand the conditions under which changes in the cellular parameters underlying learning are induced, maintained, and expressed. It has long been postulated that learning involves modifications of the efficacy or weights of the existing synaptic connections. Such experience-dependent modifications of synaptic efficacy would be expressed as an alteration in the relationship between activity in presynaptic and postsynaptic neurons and thus in the response of the network. Alternatively, the response properties of a network may be changed by experience-dependent modifications of the parameters that determine the threshold, pattern, or frequency of action potentials generated in the postsynaptic neuron in response to a given presynaptic input (i.e., the transfer or activation function of the postsynaptic cell). A third possibility is that learning involves the growth of entirely new synaptic connections and thus modifies the architecture of the network. Finally, learning may involve the nonlinear dynamic properties of individual neurons or networks such that learning is associated with shifting from one stable attractor to another. It is likely that learning involves aspects of all of these forms of plasticity.

Of the various mechanisms that have been proposed, the most extensive analyses have been done on learning rules that involve changes of synaptic weights. This chapter reviews these analyses of activity-dependent modifications of synaptic efficacy of the type believed to be involved in learning and describes several mathematical expressions that reflect some of the features of these forms of synaptic plasticity. Although many diverse rules govern the induction of synaptic plasticity, it is useful to distinguish two basic categories, nonassociative and associative. Nonassociative rules depend on one or more state variables, such as the state of activity in either the presynaptic or postsynaptic

neuron, whereas associative rules depend on the interaction of two or more state variables, such as the state of activity in the presynaptic *and* postsynaptic neurons. Moreover, associative rules generally require a specific temporal relationship between the two state variables.

NONASSOCIATIVE SYNAPTIC MODIFICATIONS

Nonassociative rules that govern the induction of synaptic plasticity are summarized in figure 4.1. There are two types: homosynaptic and heterosynaptic. Homosynaptic rules (figure 4.1A) are defined as rules in which changes to the weight of a given synapse are determined by the levels of activity in either the presynaptic or postsynaptic neurons of the given synapse (figure 4.1A, rules 1 and 2, respectively). The changes can be expressed as increases (figure 4.1A, upward arrows) or decreases (figure 4.1A, downward arrows) in synaptic strength. For heterosynaptic rules, changes in the weight of a given synapse are induced by either the levels of activity in nearby synapses (figure 4.1B, rule 3), or by the activity of modulatory systems (figure 4.1B, rule 4). Thus, homosynaptic rules depend strictly on local information or state variables (i.e., presynaptic or postsynaptic activity), whereas heterosynaptic rules depend on quasi-local (i.e., activity in nearby synapses) or global (i.e., activity in modulatory systems) state variables.

Homosynaptic Modifications

Consider first a simple synaptic arrangement in which a postsynaptic neuron, B , receives an excitatory input from a single presynaptic neuron, A (figure 4.1A). The synaptic input to cell B from cell A has a strength or weight of $w_{B,A}$. As discussed in chapter 3, the postsynaptic activity y_B is taken to be some nonlinear function f of presynaptic activity y_A and the synaptic weight. Thus, for the single synapse in figure 4.1A, the postsynaptic activity at a given time t is expressed as:

$$y_B(t) = f(w_{B,A}(t), y_A(t)). \quad (4.1)$$

During learning, the synaptic weight changes as a result of previous activity, thus:

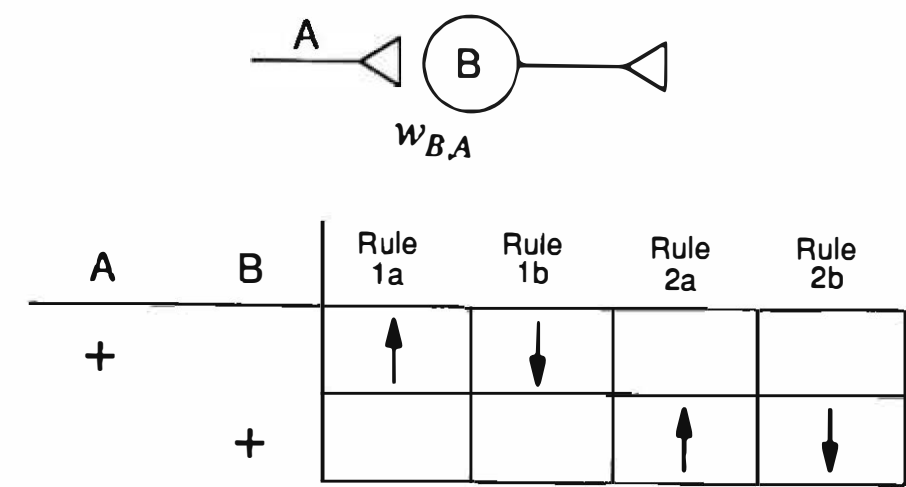
$$w_{B,A}(t+1) = w_{B,A}(t) + \Delta w_{B,A}(t) \quad (4.2)$$

where $t = 0, 1, 2, \dots$ measures discrete time and $\Delta w_{B,A}$ is the activity-dependent change in synaptic weight. A nonassociative learning rule, which is attributed to Eccles (1953), states that changes in synaptic weight depend on activity in the presynaptic cell. The most general mathematical expression that captures this notion is:

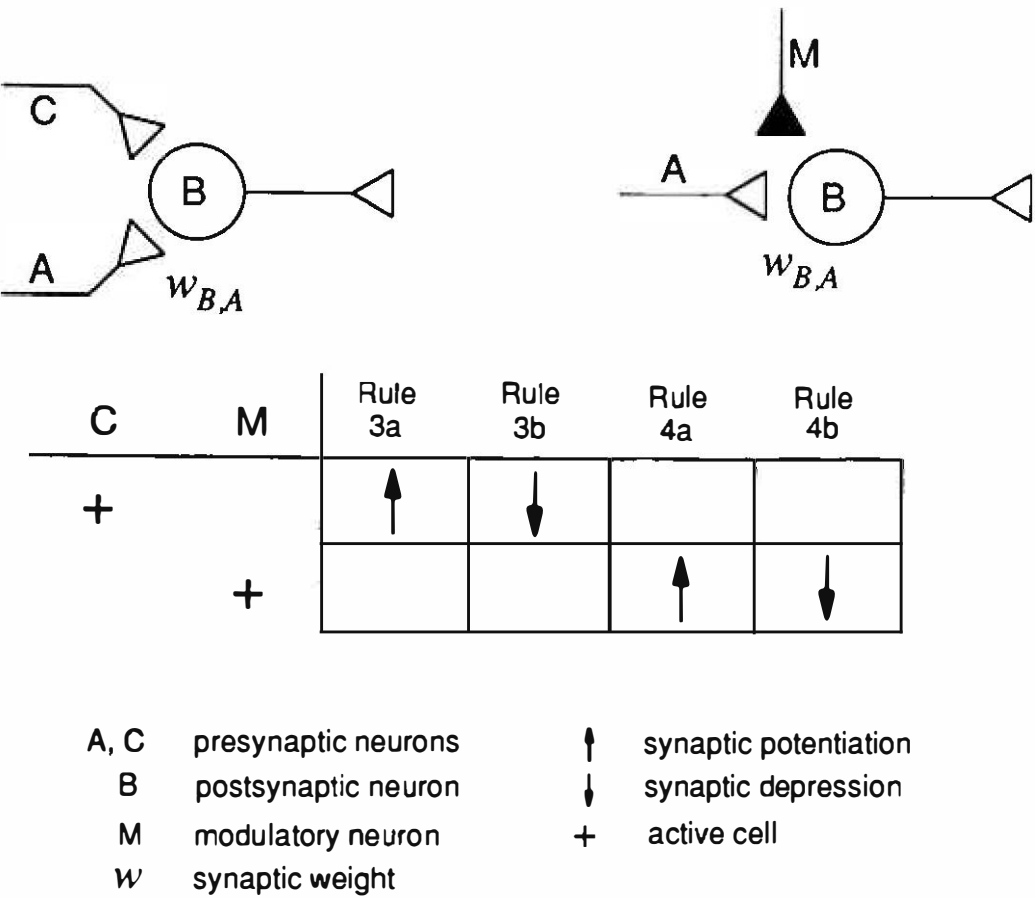
$$\Delta w_{B,A}(t) = F(y_A(t)) \quad (4.3)$$

which relates the change in synaptic weight to a function F of activity in the presynaptic neuron. Many particular learning rules fall under this

A. Nonassociative Rules for Homosynaptic Modifications of Synaptic Weights



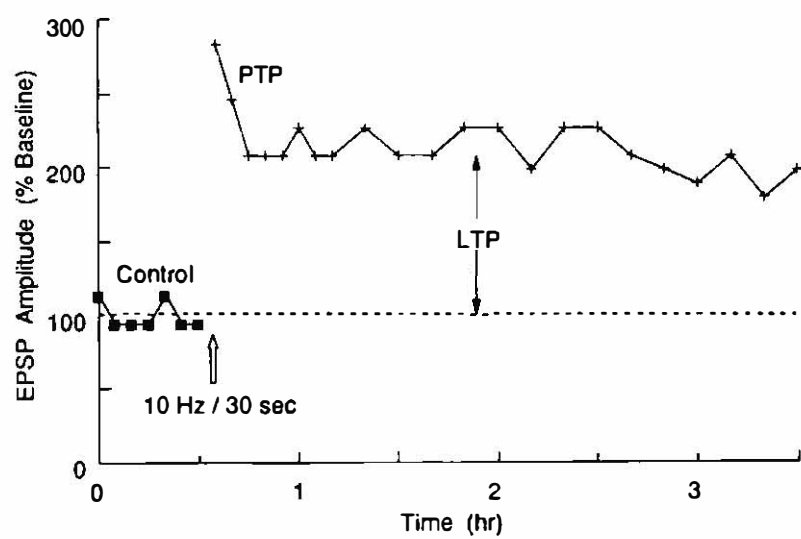
B. Nonassociative Rules for Heterosynaptic Modifications of Synaptic Weights



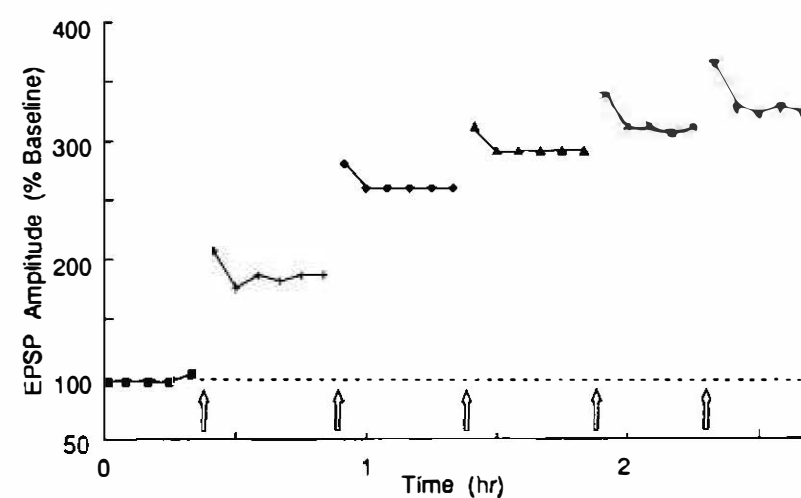
Rules for Nonassociative Synaptic Modifications

Figure 4.1 Nonassociative rules depend on one or more state variables, such as the state of activity in either the presynaptic or postsynaptic neuron. There are two classes of nonassociative rules: homosynaptic and heterosynaptic. (A) Homosynaptic rules are defined as rules in which changes to the weight of a given synapse are determined by the levels of activity in either the presynaptic (rule 1) or postsynaptic (rule 2) neuron of the synapse. The changes may be expressed as synaptic potentiation (upward arrows) or as synaptic depression (downward arrows). (B) Heterosynaptic rules are defined as rules in which changes to the weight of a given synapse are determined by either the levels of activity in nearby synapses (rule 3), or by the activity in a modulatory neuron (rule 4).

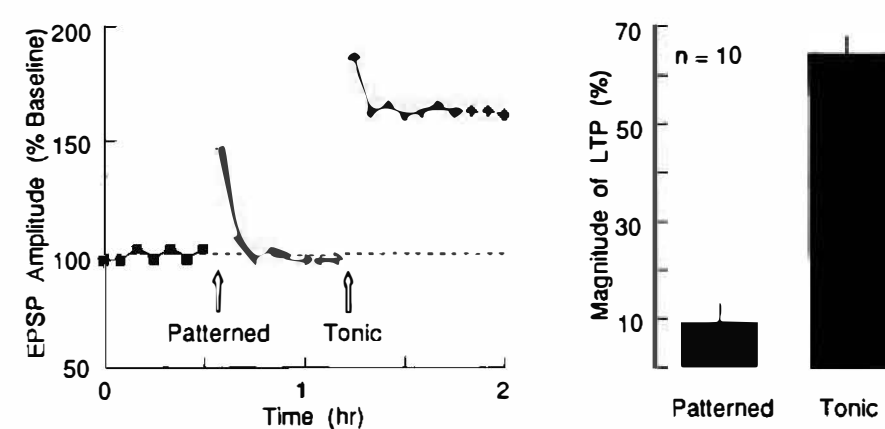
A. Long-Term Potentiation (LTP)



B. Saturation of the Induction of LTP



C. Optimal Pattern of Stimulation for Induction of LTP



Long-Term Potentiation in Opener-Neuromuscular Synapses of Crayfish

Figure 4.2 LTP, an enduring enhancement of synaptic efficacy, can be induced by a brief conditioning stimulus. (A) LTP of excitatory postsynaptic potentials (EPSPs) in opener-neuromuscular synapses of the crayfish. During the control period, EPSPs were elicited at a low frequency (0.3 Hz) that did not induce changes in the synapse. LTP was induced by stimulating the presynaptic neuron at 10 Hz for 30 s (arrow). The conditioning stimulus induced an enduring enhancement of synaptic efficacy that persisted for the duration of the experiment. (B) The LTP of EPSPs reaches an asymptote with repeated conditioning stimuli. Synaptic efficacy was monitored by eliciting EPSPs at a low frequency (0.3 Hz). Synaptic potentiation was induced by repeated conditioning stimuli of 10 Hz for 30 s (arrows). The induction of LTP appears to saturate with repeated conditioning stimuli. (C) A tonic stimulus induces LTP whereas a patterned stimulus does not. Both condition-

general form, depending on the properties of F . For example, F may be a positive function (rule 1a), in which case presynaptic activity would induce a potentiation of synaptic efficacy, or F may be a negative function (rule 1b), in which case presynaptic activity would induce a depression of synaptic efficacy.

Some forms of long-term potentiation (LTP) are examples of biological synaptic plasticity that comply with rule 1 and that provide some insights to the general features of $F(y_A)$. LTP is an enduring enhancement of synaptic efficacy, often persisting for many hours, that can be induced by a brief (several seconds) conditioning stimulus. This great asymmetry between the duration of the conditioning stimulus and the duration of the subsequent synaptic change provides an operational definition of LTP. LTP occurs in numerous synapses of the central and peripheral nervous systems of vertebrates and invertebrates, and there are likely to be many different forms of LTP. This conclusion is based, in part, on the observation that different rules have been found to govern the induction of LTP at various synapses (see below).

One example of LTP, which can be described by rule 1a, is the LTP of excitatory postsynaptic potentials (EPSPs) that has been observed in opener-neuromuscular synapses of the crayfish (Baxter et al. 1985; Bittner 1989; Keenan et al. 1987). The opener-neuromuscular synapse closely approximates the two key features of the simple, linear synaptic arrangement described in figure 4.1A and equation 4.1. First, the postsynaptic cells receive synaptic input from a single excitatory neuron. Second, the postsynaptic cells lack any voltage-gated ionic conductances that would introduce nonlinearities. Thus, the presynaptic release of the excitatory transmitter glutamate linearly depolarizes the postsynaptic cells over a wide range of membrane potentials, and the amplitudes of the EPSPs reflect the level of presynaptic activity and the efficacy of the synapse (equation 4.1). Figure 4.2A illustrates some experimental results. The efficacy of the synapse was tested by stimulating the presynaptic neuron at a low frequency (0.3 Hz) that did not induce changes in the synapse during the 30-min control period. In this experiment, potentiation was induced by eliciting action potentials in the presynaptic neuron at 10 Hz for 30 s. The potentiation had two kinetically distinct components. The first component decays rapidly with a time course corresponding to that of posttetanic potentiation (PTP; for a review of PTP see Zucker 1989). PTP is followed by a much longer enhancement of synaptic efficacy that persists for the duration of the experiment, in

ing stimuli delivered 300 presynaptic action potentials over 30 s. In the tonic stimulus, the 300 presynaptic spikes are elicited at a constant frequency of 10 Hz, whereas in the patterned stimulus, the 300 presynaptic spikes are arranged as brief (2 to 5 spikes), high-frequency (100 Hz) bursts that were evenly distributed throughout the 30 s. The results of ten experiments are summarized. The magnitude of LTP (the percent increase above the control value) was measured 30 min after the conditioning stimulus. (Modified from Keenan et al. 1987.)

some cases up to 8 hr. This latter component satisfies the operational definition of LTP.

Several lines of evidence suggest that the induction of LTP in the opener-neuromuscular synapse is nonassociative. First, there are no other excitatory synaptic inputs to the postsynaptic cells, and there is no modulatory system in the isolated claw preparation. Second, it is unlikely that postsynaptic depolarizations are necessary for LTP in this synapse since the postsynaptic cells do not fire action potentials and neither the test stimuli nor the conditioning stimulus significantly depolarize the postsynaptic cell. Third, a quantal analysis indicated that LTP could be accounted for by an increase in the presynaptic release of transmitter (Baxter et al. 1985). Thus, LTP in the opener-neuromuscular synapse is believed to be an example of a nonassociative, homosynaptic form of LTP that is induced by presynaptic activity (rule 1a). A similar rule appears to govern, or at least contribute to, the induction of LTP in several other excitatory synapses, including sensorimotor synapses of *Aplysia* (Walters and Byrne 1985), mammalian sympathetic ganglia (Alonso-deFlorida et al. 1991; Bachoo and Polosa 1991; Briggs and McAfee 1988), bullfrog sympathetic ganglia (Koyano et al. 1985), and the mossy fiber inputs to pyramidal neurons in the CA3 region of the hippocampus (Zalutsky and Nicholl 1990; see however, Jaffe and Johnston 1990).

Perhaps the simplest form of F (rule 1a) is a simple product:

$$\Delta w_{B,A}(t) = \varepsilon y_A(t) \quad (4.4)$$

where ε is a positive constant that determines the rate of increase in synaptic strength (often referred to as the learning rate) and y_A is some measure of presynaptic activity. The most common interpretation of y_A is that this state variable represents the frequency of presynaptic spike activity, which is averaged over several seconds prior to time t . The two key assumptions in equation 4.4 are that $\Delta w_{B,A}$ is a monotonically increasing function of presynaptic activity and that the proper measure for governing the induction of LTP is the average frequency of presynaptic activity. However, our current understanding of nonassociative LTP does not reflect these assumptions.

The first problem with equation 4.4 is that repeated conditioning stimuli do not induce ever-increasing synaptic potentiation. Rather, as illustrated in figure 4.2B, potentiation asymptotically approaches a maximum. The five upward arrows in figure 4.2B indicate the occurrences of conditioning stimuli (presynaptic stimulation at 10 Hz for 30 s). The accumulation of LTP with repeated conditioning stimuli resembles the negatively accelerating acquisition curves of the sort found in the Rescorla and Wagner (1972; see also Donegan et al. 1989) learning equations. Combining equations 4.2 and 4.4, and incorporating an expression that places asymptotic limits on synaptic strength yields:

$$w_{B,A}(t + 1) = w_{B,A}(t) + [\varepsilon y_A(t) (\lambda_{\max} - w_{B,A}(t))] \quad (4.5)$$

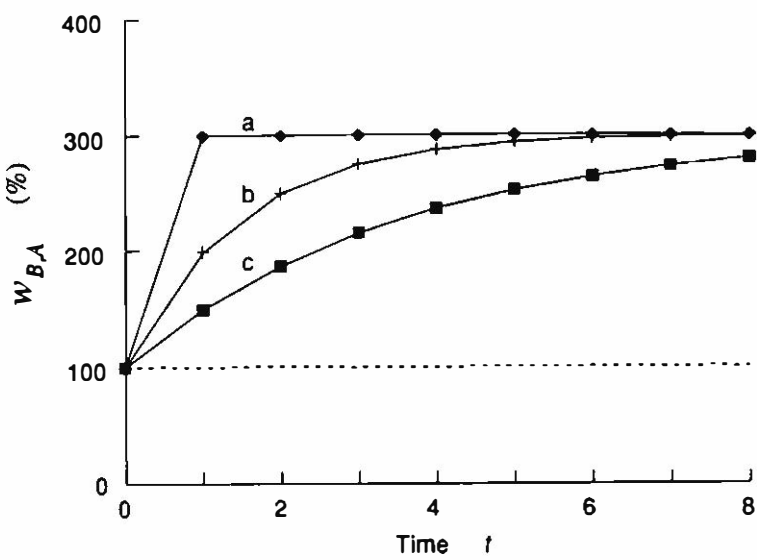
where $\lambda_{\max}$ is the maximum possible synaptic strength. Including $\lambda_{\max}$ as a limit for $\Delta w_{B,A}$ introduces an additional problem in that synaptic strength will inevitably saturate. Thus, some mechanism must be included to allow for the decrement of synaptic plasticity (see below).

It is possible to estimate values for $\lambda_{\max}$ and ε from current experimental data. The maximum magnitude of LTP has been found in a number of different vertebrate and invertebrate preparations to range between a 50% to 600% increase above the control value of $w_{B,A}$ (i.e., the value of $w_{B,A}$ at time $t = 0$). Thus, the value of $\lambda_{\max}$ should be no more than about six to ten times the value of $w_{B,A}(t = 0)$. It should be noted that $\lambda_{\max}$ may not represent the full dynamic range of the synapse. For example, in opener-muscular synapses, short-term facilitation can enhance synaptic transmission by several orders of magnitude (Bittner 1968, 1989; Zucker 1989), whereas the maximum magnitude of LTP is only about a three or four fold increase (figure 4.2B). LTP appears to approach its asymptotic value following only a few conditioning stimuli. Thus, the value of ε should be between about 0.01% to 0.5% of y_A . Figure 4.3A illustrates three numerical examples of the growth of $w_{B,A}$ that is predicted by equation 4.5 at three different stimulus strengths.

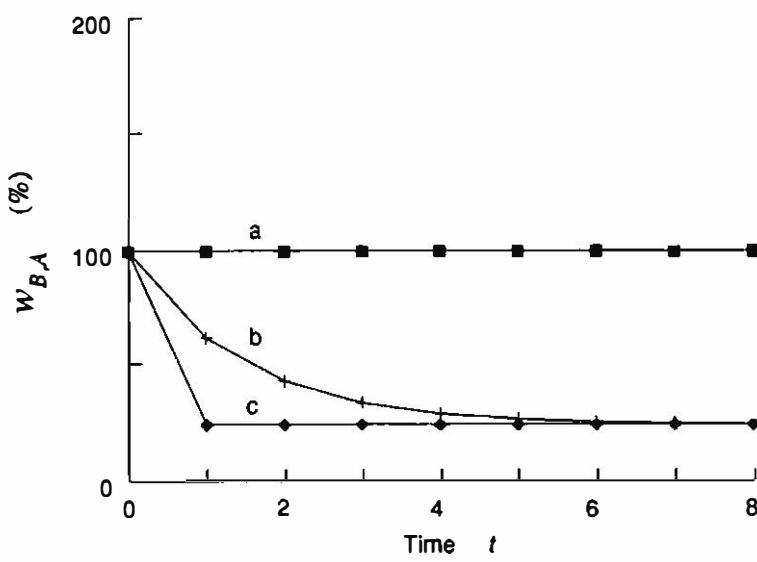
A second problem with equation 4.4 is that $\Delta w_{B,A}$ will always tend to increase because ongoing presynaptic activity will contribute to the time-averaged value of y_A . One solution to this problem is to compute $\bar{y}$, a measure of the sustained or background level of activity in the presynaptic neuron averaged over a much longer time span than y_A , and to use a function of $(y_A - \bar{y}_A)$ as the presynaptic expression in equation 4.5. Thus, potentiation would occur only when $y_A > \bar{y}_A$, and no change will occur when $y_A \leq \bar{y}_A$.

A third problem with equation 4.4 is that the average of presynaptic spike activity may not be the proper measure with which to control the induction of synaptic potentiation. For example, figure 4.2C illustrates an experiment that examines the ability of two conditioning stimuli to potentiate opener-neuromuscular synapses. The two conditioning stimuli are termed “tonic” and “patterned.” Both conditioning stimuli have a duration of 30 s and elicit 300 presynaptic action potentials, i.e., an average presynaptic spike activity of 10 Hz for 30 s. In the tonic stimulus, the 300 spikes are elicited at a constant frequency of 10 Hz, whereas in the patterned stimulus, the 300 spikes are arranged as brief (2 to 5 spikes), high-frequency (100 Hz) bursts that were evenly distributed throughout the 30-s period. The patterned conditioning stimulus failed to induce potentiation, whereas the tonic conditioning stimulus induced potentiation. These results suggest that the induction of potentiation is governed by specific features of the conditioning stimulus. In addition, other experimental data suggest that various aspects of synaptic potentiation, such as the magnitude and duration, can be differentially con-

A. Equation 4.5



B. Equation 4.7



Effects of Different Stimulus Strengths on the Solutions of Equations 4.5 and 4.7

Figure 4.3 Equations 4.5 and 4.7 predict synaptic potentiation and depression, respectively, and place asymptotic limits on changes that can occur in synaptic efficacy. (A) As the presynaptic stimulus strength in equation 4.5 is increased from $y_A = 25$ (trace c) to $y_A = 50$ (trace b) to $y_A = 100$ (trace a), potentiation of $w_{B,A}$ reaches its asymptotic level with a smaller number of trials ($\epsilon = 0.01$, $\lambda_{\max} = 3$, and $w_{B,A}(t = 0) = 1$ in all three traces). For simplicity, each of these terms is dimensionless. Note that the strength of the conditioning stimulus y_A affects only the rate of growth, not the maximum level of synaptic potentiation. (B) Because of the inverse relationship between synaptic depression and the frequency of presynaptic activity in equation 4.7, a very low stimulus strength ($y_A = 0.01$, trace c) rapidly induces maximal depression of $w_{B,A}$ ($\epsilon = 0.01$, $\lambda_{\min} = 0.25$, and $w_{B,A}(t = 0) = 1$ in all three traces). The rate of depression is slower as the stimulus strength is increased slightly to $y_A = 0.02$ (trace b). In contrast, more intense stimuli ($y_A = 3$, trace a) induce no observable depression.

trolled by parameters of the conditioning stimulus (Alonso-deFlorida et al. 1991; see also Grover and Teyler 1990). Thus, it appears that while equation 4.5 reflects some of the simpler features of synaptic potentiation, such as asymptotic limits to synaptic potentiation, it does not accurately reflect some of the more subtle properties of potentiation, such as the stimulus-specificity that governs induction. An alternative approach is to develop dynamic models of synaptic transmission that incorporate mathematical formalisms of the underlying cellular mechanisms. Such biologically realistic models are likely to express more precisely the features of synaptic transmission and plasticity. One such model, activity-dependent neuromodulation of synaptic efficacy, is described later in this chapter.

The stimulus specificity that is depicted in figure 4.2C illustrates a potentially intriguing aspect of information processing and storage in a single synapse. Namely, information processing and information storage can be controlled independently in a single synapse by different patterns of presynaptic activity. For example, in the opener-neuromuscular synapse, brief bursts of high-frequency presynaptic activity are the optimum stimulus for eliciting a behavioral postsynaptic response, consisting of significant postsynaptic depolarization and muscle contraction (Bittner 1968; Smith 1975). This type of stimulus, however, does not appear to induce LTP. In contrast, brief trains of low-frequency presynaptic activity do not elicit a muscle contraction, but do induce LTP. Thus, long-term changes in synaptic efficacy may be induced without significant activity in the postsynaptic cell, and conversely the postsynaptic cell can be activated without inducing long-term changes in the synapse. Moreover, once induced in opener-neuromuscular synapses, LTP was found to have a multiplicative (not additive) effect on the short-term synaptic plasticity (Keenan et al. 1987). Thus, LTP functions to increase the “gain” of information processing throughout the entire dynamic range of the synapse.

The inverse of LTP is long-term depression (LTD), a type of synaptic plasticity in which the efficacy of a synapse is persistently reduced. A form of nonassociative, homosynaptic depression that is induced by presynaptic activity (rule 1b) has been observed at several excitatory synapses, including the neuromuscular synapses in crayfish (Pahapill et al. 1987; Zucker and Bruner 1977), sensorimotor synapses in *Aplysia* (Byrne 1982; Castellucci et al. 1970), and inputs to the CA1 pyramidal cells of the rat hippocampus (Dudek and Bear 1992; Dunwiddie and Lynch 1978; Mulkey and Malenka 1992). Recent results indicate that the postsynaptic cell may play a role in the induction of synaptic depression in CA1 neurons (Dudek and Bear 1992; Mulkey and Malenka 1992). Thus, synaptic depression in CA1 neurons may turn out to be associative rather than nonassociative.

At many synapses, the magnitude of depression has been found to be inversely related, within a range of activity, to the frequency of presyn-

aptic activity. Thus, low frequencies of presynaptic stimulation (e.g., 0.01 Hz) induce more depression than higher frequencies of stimulation (e.g., > 1 Hz). Two expressions, which are fashioned after equations 4.4 and 4.5, can be used to describe synaptic depression. The first is a simple product in which the learning rate ε is now negative:

$$\Delta w_{B,A}(t) = \varepsilon (y_A(t))^{-1} \quad (4.6)$$

In this equation and the next, y_A is restricted to a specific range of values that produces depression. This excludes, for example, $y_A = 0$; in the absence of presynaptic activity there should be no change in $w_{B,A}$. In addition, although this equation produces decreases in $w_{B,A}$, it is again not a biologically plausible expression in that it can lead to negative values for synaptic weights. In an alternate formulation, the learning rate ε can remain a positive value, but the asymptote λ for changes in synaptic strength is made to be less than the initial value of $w_{B,A}$ ($w_{B,A}$ at $t = 0$):

$$w_{B,A}(t + 1) = w_{B,A}(t) + [\varepsilon (y_A(t))^{-1} (\lambda_{\min} - w_{B,A}(t))] \quad (4.7)$$

where $\lambda_{\min}$ is the minimum possible synaptic strength and values for $\lambda_{\min}$ are restricted to the range of $0 < \lambda_{\min} < w_{B,A}(t = 0)$. Figure 4.3B illustrates three numerical examples of the depression of $w_{B,A}$ that is predicted by equation 4.7.

Although the values of $\lambda_{\max}$ and $\lambda_{\min}$ in equations 4.5 and 4.7 are generally assumed to be constants, recent experimental evidence suggests that their values may depend on the state of the synapse. For example, a low-frequency conditioning stimulus (e.g., 100 pulses at 1 Hz) applied to Schaeffer collateral/commissural (Sch/comm) projections to CA1 pyramidal cells in the hippocampus selectively depressed potentiated synapses but had no effect on control synapses (Barrionuevo et al. 1980; Staubli and Lynch 1990). This result implies that the type of response (potentiation vs. no change vs. depression) induced by a given conditioning stimulus may depend on the strength of the synapse prior to the conditioning stimulus. Thus, a critical experimental issue is the determination of the intrinsic and/or extrinsic factors that control the state-dependent responses of synapses. An intriguing possibility is that the state-dependent responses of synapses may be influenced by modulatory transmitters. Thus, the plasticity at a given synapse could, in principle, be linked to more global factors that are associated with motivational or developmental states.

Shimbel (1950) proposed a learning rule in which changes in synaptic weight depend on postsynaptic activity (figure 4.1A, Rule 2). A general expression of this learning rule is:

$$\Delta w_{B,A}(t) = F(y_B(t)) \quad (4.8)$$

which relates the change in synaptic weight to a function F of postsynaptic activity y_B . Several examples of biological synaptic plasticity (see

below) comply with rule 2, and indicate postsynaptic activity can induce either synaptic potentiation (rule 2a) or depression (rule 2b).

A form of nonassociative, homosynaptic potentiation that is induced by postsynaptic activity (rule 2a) has been observed at two excitatory synapses: bullfrog sympathetic ganglion (Kumamoto and Kuba 1983) and stellate cells in layer II of the entorhinal cortex of the guinea pig (Alonso et al. 1990). In the bullfrog sympathetic ganglion, potentiation was induced by antidromic stimulation (20 Hz for 5 s) of the postganglionic nerve trunk. This postsynaptic electrical activity, which was not paired with presynaptic activity, led to a doubling of the amplitudes of the EPSPs. In the entorhinal cortex, potentiation of the projection from the piriform cortex to stellate cells in layer II was induced by depolarizing the postsynaptic cell. This unpaired postsynaptic conditioning stimulus led to an approximate doubling of synaptic efficacy.

A mathematical expression for postsynaptically induced potentiation, can be fashioned after equation 4.5 by replacing the presynaptic factor y_A with a term y_B that represents some measure of postsynaptic activity, such as the frequency of spike activity averaged over several seconds prior to time t . Experimental data suggest that the value of $\lambda_{\max}$ should be no more than about three times the control value of $w_{B,A}$. As discussed above, such a simple mathematical expression may not reflect some of the more subtle properties of postsynaptically induced synaptic plasticity. For example, the average of postsynaptic spike activity may not be the pertinent state variable. In the entorhinal cortex, potentiation was induced by subthreshold membrane depolarizations; i.e., there were no postsynaptic spikes. Moreover, only rhythmic postsynaptic depolarizations that resembled the intrinsic membrane potential oscillations that are present in the postsynaptic cells were effective in inducing potentiation. Thus, the postsynaptic induction of synaptic potentiation appears to be governed by specific features of the conditioning stimulus.

A form of nonassociative, homosynaptic depression that is induced by postsynaptic activity (rule 2b) has been observed in Sch/comm projections to CA1 pyramidal neurons in the hippocampus (Pockett et al. 1990; see also Dunwiddie and Lynch 1978). Depression of excitatory synaptic transmission was induced by antidromic stimulation (e.g., six bursts of 50 stimuli at 100 Hz, delivered at 10-s intervals) of the axons of CA1 pyramidal neurons. This postsynaptic electrical activity, which was not paired with presynaptic activity, decreased the amplitudes of EPSPs to as little as 14% of their control values. A function, which approximately describes postsynaptic induced depression, can be fashioned after equation 4.5 first by substituting y_B (postsynaptic activity) for the presynaptic term and by restricting the values of $\lambda_{\min}$ to a range of $0 < \lambda_{\min} < w_{B,A}(t = 0)$ (i.e., $\lambda_{\min}$ must be less than the control value for synaptic strength). Experimental data suggest that the lower limit of $\lambda_{\min}$ should be about 0.14.

Heterosynaptic Modification

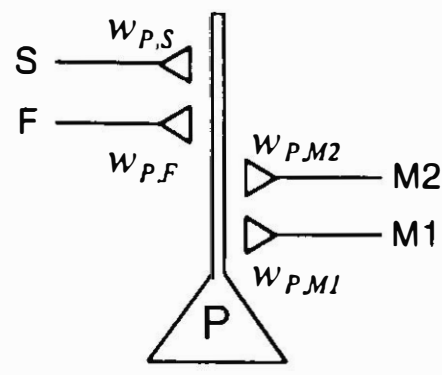
Consider a common synaptic arrangement in which a postsynaptic neuron, B , receives converging excitatory inputs from several presynaptic neurons, such as presynaptic neurons A and C in figure 4.1B. Experimental data suggest a learning rule in which changes in the synaptic weight of cell A can be a function of the activity in the neighboring presynaptic neuron C :

$$\Delta w_{B,A}(t) = F(y_C(t)) \quad (4.9)$$

where y_C is some measure of activity in neuron C . Activity in neuron C can either potentiate (rule 3a) or depress (rule 3b) the synaptic strength of the nonactive cell, neuron A . A mathematical expression for this form of heterosynaptic plasticity (rule 3) can be fashioned from equation 4.5 by replacing y_A with y_C . As before, this expression will represent synaptic potentiation if the value of $\lambda_{\max}$ is greater than $w_{B,A}(t=0)$ (the control value of synaptic strength) and depression if the value of $\lambda_{\min}$ is less than $w_{B,A}(t=0)$.

A form of nonassociative, heterosynaptic potentiation that can be induced in a set of nonactive synapses by activity in neighboring synapses (rule 3a) has been observed at several excitatory synapses, including inputs to the stellate ganglion of the cat (Bachoo and Polosa 1991) and certain classes of inputs to pyramidal cells in the CA3 region of the hippocampus (Bradler and Barrionuevo 1989, 1990). In the stellate ganglion, synaptic potentiation was examined by stimulating two distinct groups of preganglionic axons, the third and fourth thoracic white rami (T3WR and T4WR, respectively). A conditioning stimulus (e.g., 40 Hz for 5 s) applied to T4WR, while T3WR was unstimulated, induced heterosynaptic potentiation in the T3WR pathway and increased the T3WR response by an average of 600% above the control values. If the conditioning stimulus was applied to the test pathway, T3WR, homosynaptic potentiation was induced in T3WR. If the conditioning stimulus was applied to the postganglionic axons, no changes in synaptic efficacy were induced. Thus, the heterosynaptic potentiation in this preparation was not a consequence of postsynaptic electrical activity and these synapses can express both homosynaptic and heterosynaptic potentiation. Currently, little is known about the ways in which these two forms of synaptic plasticity interact within a single synapse.

A form of nonassociative, heterosynaptic depression that can be induced in a set of nonactive synapses by activity in neighboring synapses (rule 3b) has been observed in several excitatory synapses in the hippocampus, including inputs to: pyramidal cells in the CA1 (Dunwiddie and Lynch 1978; Lynch et al. 1977), pyramidal cells in the CA3 (Bradler and Barrionuevo 1989, 1990), and granule cells in the dentate gyrus (Abraham and Goddard 1983; Abraham et al. 1985; Levy 1985; Levy and Steward 1979, 1983; White et al. 1988, 1990). For example, in the CA3



M1	S	F	$w_{P,M1}$	$w_{P,M2}$	$w_{P,S}$	$w_{P,F}$
+			Homosynaptic Potentiation	Heterosynaptic Depression	Heterosynaptic Potentiation	Heterosynaptic Potentiation
	+		Heterosynaptic Depression		Homosynaptic Potentiation	
		+	Heterosynaptic Depression		Heterosynaptic Depression	Homosynaptic Potentiation

M1, M2

S

P

mossy fiber pathways

Schaffer collateral / commissural pathways

CA3 pyramidal neuron

F

fimbrial pathway

w

synaptic weight

+

tetanized pathway

Synaptic Plasticity Observed at Inputs to Hippocampal CA3 Pyramidal Neurons

Figure 4.4 Heterosynaptic plasticity at inputs to a hippocampal CA3 pyramidal neuron (P) was examined by stimulating four separate groups of presynaptic fibers: Schaeffer collateral/commissural (S), fimbrial (F), and two distinct groups of mossy fibers (M1, M2). Stimulation (+) of a subset of the mossy fiber input, M1, induced heterosynaptic depression in the nonstimulated mossy fiber input, M2, and heterosynaptic potentiation in the fimbrial and Schaeffer pathways. Stimulation of either the Schaeffer collateral/commissural or fimbrial inputs also induced heterosynaptic depression in the nonstimulated mossy fiber pathway. (Data summarized from Bradler and Barrionuevo 1990.)

region of the hippocampus, heterosynaptic depression was examined by stimulating up to four separate groups of presynaptic fibers: Schaeffer collateral/commissural (S), fimbrial (F), and two distinct groups of mossy fibers (M1, M2) (figure 4.4). A conditioning stimulus (five 6-pulse 100 Hz trains delivered at 0.02 Hz) applied to a subset of the mossy fiber inputs (M1) induced heterosynaptic depression in a nontetanized subset of mossy fibers, M2, whereas heterosynaptic potentiation was induced in the fimbrial and Schaeffer pathways. Heterosynaptic depression of the mossy fiber pathways also could be induced by conditioning stimuli applied to either the Schaeffer or fimbrial pathways. Only the mossy fiber pathway was capable of inducing heterosynaptic potentiation and this heterosynaptic potentiation was restricted to non-mossy fiber pathways.

A feature of heterosynaptic plasticity not captured by rules similar to equation 4.5 is the spatial limitation that may apply to interactions among synapses. For example, White et al. (1988, 1990) found in the dentate gyrus that heterosynaptic depression could be induced only between

83

Chapter 4

Learning in the hippocampus

populations of synapses that converged onto overlapping dendritic regions. Indeed, the degree of heterosynaptic interactions, as measured by the magnitude of depression, was correlated with the extent of spatial overlap (i.e., the strength of heterosynaptic interactions between pathways was directly related to the proximity of the interacting synapses). Thus, mathematical expressions of heterosynaptic interactions should incorporate some description of the anatomical features of the dendritic tree and of the spatial arrangements of different populations of synapses on the dendrites. The spatial limitations of heterosynaptic interactions also suggest an intriguing aspect to single-cell information processing and storage—namely, different regions of a cell can process and store information independently. To capture this feature, single lumped parameter models of the postsynaptic cell are not appropriate.

As illustrated in figure 4.1B, the heterosynaptic interactions between neurons can be mediated by modulatory transmitters; that is, changes in synaptic weight are a function of the activity in a modulatory neuron or system, y_M :

$$\Delta w_{B,A}(t) = F(y_M(t)). \quad (4.10)$$

These modulatory neurons can either potentiate (rule 4a) or depress (rule 4b) the synaptic efficacy of neuron A . A particularly well-characterized example of such heterosynaptic interactions is the modulation of the sensory neurons that mediate defensive withdrawal reflexes in *Aplysia* (for reviews, see Byrne 1987; Byrne and Crow 1991; Byrne et al. 1991a, c; 1993; Kandel and Schwartz 1982). The properties of these sensory neurons can be affected by several modulatory transmitters, some of which induce heterosynaptic potentiation while others induce heterosynaptic depression. For example, serotonin (5-HT) can induce nonassociative heterosynaptic potentiation in sensory neuron synapses. Binding of serotonin to its receptors activates several second messenger/kinase systems in the sensory neurons, two of which are the cAMP-dependent protein kinase A system and the Ca^{2+} /phospholipid-dependent protein kinase C system (Bernier et al. 1982; Ocorr and Byrne 1985; Sacktor and Schwartz 1990). These second messenger/kinase systems, in turn, modulate the properties of ion channels, which results in increased neuronal excitability, increased spike duration, increased Ca^{2+} influx, and hence increased transmitter release from sensory neurons (Baxter and Byrne 1990; Klein and Kandel 1980; Sugita et al. 1992). In addition, these second messenger/kinase systems seem to directly modulate aspects of the secretory machinery, which in turn enhances transmitter release (Braha et al. 1990; Gingrich and Byrne 1984, 1985, 1987; Gingrich et al. 1988; Hochner et al. 1986; Pieroni and Byrne 1992).

In contrast, the neuropeptide FMRFamide can induce nonassociative heterosynaptic depression in sensory neuron synapses. FMRFamide activates the arachidonic acid second messenger pathway (Piomelli et al. 1987), and in general, the effects of FMRFamide are the opposite of

those induced by serotonin. FMRFamide modulates ion channels in such a way that neuronal excitability, spike duration, Ca^{2+} influx, and hence transmitter release, are all reduced (Belardetti and Siegelbaum 1988; Critz et al. 1991; Edmonds et al. 1990). In addition, FMRFamide seems to directly inhibit aspects of the secretory machinery, which in turn reduces transmitter release (Pieroni and Byrne 1992; see also Manson-Hing et al. 1989) and may have postsynaptic actions as well (Peter et al. 1992).

Thus, the actions of the modulatory transmitters that induce heterosynaptic modification are mediated via a complex assembly of interacting second messenger systems and enzymatic reactions that affect multiple subcellular processes. This complexity in the mechanisms that underlie heterosynaptic modulation can not be easily manifested by relatively simple mathematical expressions such as equations 4.5 or 4.10. More biologically realistic and dynamic models of the types of subcellular and molecular mechanisms that can underlie heterosynaptic modulation have been developed (Aszódi and Friedrich 1987; Aszódi et al. 1991; Buxbaum and Dudai 1989; Friedrich 1990; Gingrich and Byrne 1985, 1987; Gingrich et al. 1988; Greenberg et al. 1987; see below). These models illustrate that the molecular processes underlying neuronal functions, including the kinetics of enzyme reactions, interactions among second messenger systems, and regulation of transmitter release, play critical roles in determining the features of neuronal plasticity.

ASSOCIATIVE SYNAPTIC MODIFICATIONS

Associative modifications of synaptic efficacy depend on the temporal correlation between the activities in two neurons. Although several associative learning rules have been proposed on the basis of theoretical work, only recently have examples of associative synaptic modifications been demonstrated experimentally. Below are reviewed several examples of biological synaptic plasticity that are governed by the temporal interactions between two neuronal stimuli and that provide some insight to the formulation of biologically plausible associative learning rules.

Hebbian Rule for Synaptic Modification

Possibly the best known associative learning rule was proposed by Hebb (1949). His postulate for learning states:

When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency as one of the cells firing B is increased.

The key feature of the Hebbian rule is that increases in synaptic weight are dependent on concurrent activity in the presynaptic and postsyn-

aptic cells. This simple conjunctive mechanism for synaptic potentiation can be represented by the general expression:

$$\Delta w_{B,A}(t) = F(y_A(t), y_B(t)) , \quad (4.11)$$

which relates the change in synaptic weight to a function F of both presynaptic and postsynaptic activity.

At least one form of synaptic potentiation, which has been examined extensively in the Sch/comm projections to CA1 pyramidal cells in the hippocampus, appears to conform to the Hebbian rule. A weak presynaptic stimulus, which by itself fails to induce any change in synaptic efficacy, can induce LTP when paired with intracellular depolarizations of the postsynaptic cell (Gustafsson et al. 1987; Kelso et al. 1986; Wigström et al. 1986). This form of LTP can be induced even when Na^+ -mediated postsynaptic action potentials are blocked (Kelso et al. 1986). Evidently, it is sufficient that the postsynaptic cell to be strongly depolarized at the same time that the presynaptic cells are stimulated. The temporal specificity of this form of synaptic potentiation has been examined by varying the intervals between postsynaptic depolarization and presynaptic stimulation (Gustafsson et al. 1987; Kelso et al. 1986; see also Brown et al. 1989; Gustafsson and Wigström 1986; Hashemzadeh et al. 1991; Levy and Steward 1983). Simultaneous occurrences of presynaptic stimulation and postsynaptic depolarization are the most effective at inducing potentiation. In addition, potentiation can be induced with postsynaptic depolarizations that start several tens of milliseconds after the presynaptic activity, but not with postsynaptic depolarizations that terminate before the presynaptic activity. These results suggest that following presynaptic stimulation there remains a "trace" of previous activity during which time the synapses continue to be eligible for modification. (For other consequences of this trace, see chapter 2.)

Although neither firing postsynaptic action potentials per se nor strictly conjunctive activity may be required, potentiation of Sch/comm projections to CA1 cells is generally consistent with the Hebbian rule in that synaptic potentiation takes place only when the synapse is active in close temporal contiguity with strong postsynaptic depolarizations. A similar rule appears to govern the potentiation in excitatory projections from the entorhinal cortex to the dentate gyrus of the hippocampus (for review see Levy 1985), as well as excitatory synapses in a wide variety of brain areas (for reviews, see Teyler and DiScenna 1987; Teyler et al. 1990; Tsumoto 1990) and possibly some excitatory synapses of invertebrates (Miller et al. 1987).

There have been many mathematical formalisms suggested for the Hebbian rule (for reviews see Brown et al. 1990; Sejnowski and Tesauro 1989, 1990). The simplest implementation of the Hebbian rule is a product of the average frequencies of spike activity in the presynaptic and postsynaptic cells:

$$\Delta w_{B,A}(t) = \epsilon y_A(t) y_B(t) \quad (4.12)$$

where ε is a small, positive constant. As discussed above (see discussion of equation 4.4), there are problems with a simple product equation. For example, the product equation predicts that $\Delta w_{B,A}$ is a monotonically increasing function of conjunctive activity. This prediction is in contrast, however, to experimental observations in which potentiation of Sch/comm projections to CA1 cells saturates with repeated conditioning stimuli (Gustafsson and Wigström 1986; Larson and Lynch 1986; see also Abraham et al. 1985; Levy and Steward 1979). Although this problem can be overcome by including an upper limit for $\Delta w_{B,A}$ such as $\lambda_{\max}$ in equation 4.5, limiting $\Delta w_{B,A}$ introduces the additional problem that chance coincidences of presynaptic and postsynaptic activity will inevitably lead to saturation of the synapse.

Brown et al. (1990; see also Kairiss et al. 1988; Kohonen 1984, 1990) suggested a more flexible expression: a bilinear equation that combines the product equation for Hebbian potentiation and mechanisms for decreases in synaptic strength:

$$w_{B,A}(t+1) = w_{B,A}(t) + (\varepsilon y_A(t) y_B(t)) - (\beta y_B(t)) - (\gamma y_A(t)) - \delta \quad (4.13)$$

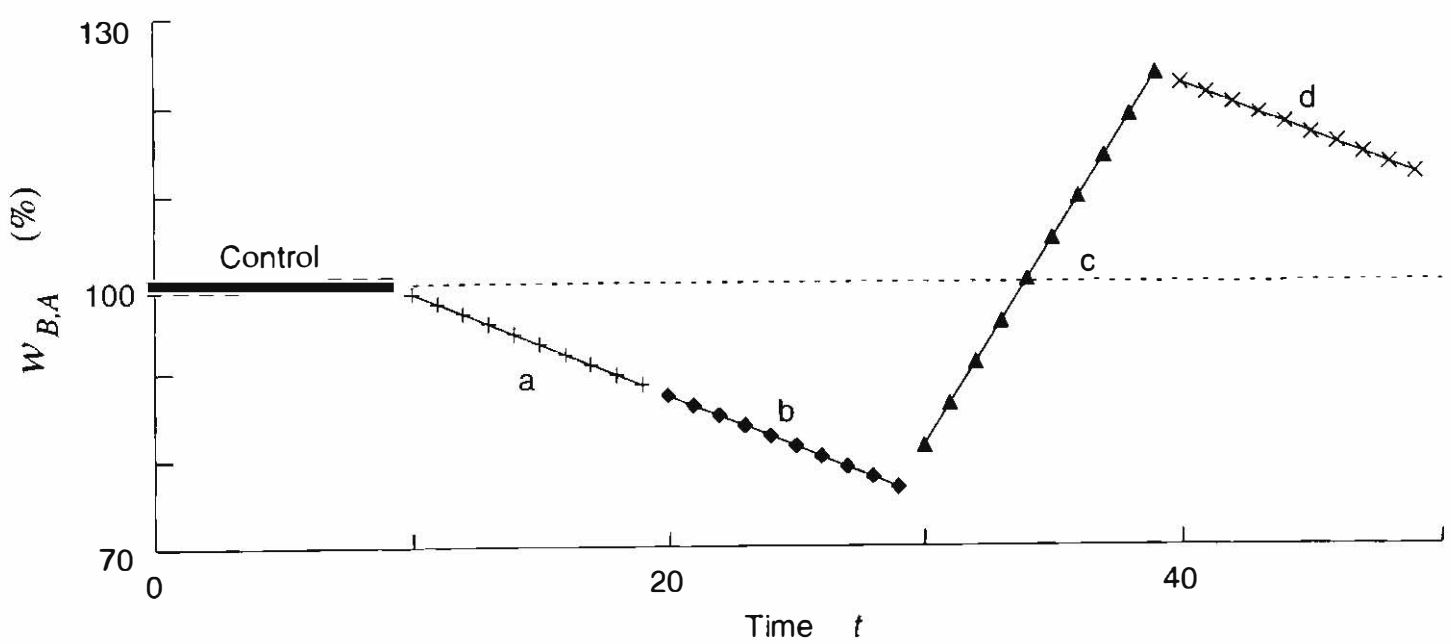
where β , γ and δ are constants in the simplest case, or functions in a more generalized form. Thus, the term $(-\beta y_B(t))$ represents a form of nonassociative synaptic depression that is induced by postsynaptic activity (Pockett et al. 1990), the term $(-\gamma y_A(t))$ represents a form of nonassociative synaptic depression that is induced by presynaptic activity (Dunwiddie and Lynch 1978), and the term $(-\delta)$ represents the passive decay of plasticity. Figure 4.5A illustrates the response of this bilinear equation (equation 4.13) to four phases of stimulation. During the first phase (phase a in figure 4.5A), only the presynaptic cell is active and $w_{B,A}$ undergoes homosynaptic depression. During the second phase (phase b), only the postsynaptic cell is active and $w_{B,A}$ again is depressed. During conjunction activity (phase c), $w_{B,A}$ potentiates as predicted by the product equation for the Hebbian rule. Finally, noncoincident activity (phase d) again results in synaptic depression.

Levy and co-workers (Levy 1985, 1989; Levy and Colbert 1991; Levy and Desmond 1985; Levy et al. 1983; Lopez et al. 1990) also have suggested a mathematical expression for the Hebbian rule that allows for reversible synaptic modifications and that incorporates asymptotic limits on both potentiation and depression of synaptic weight:

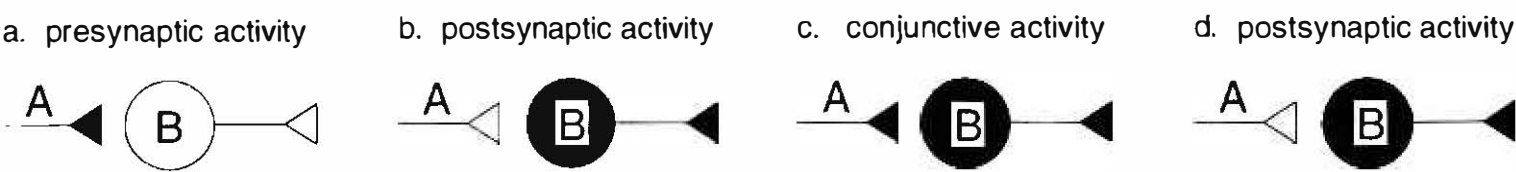
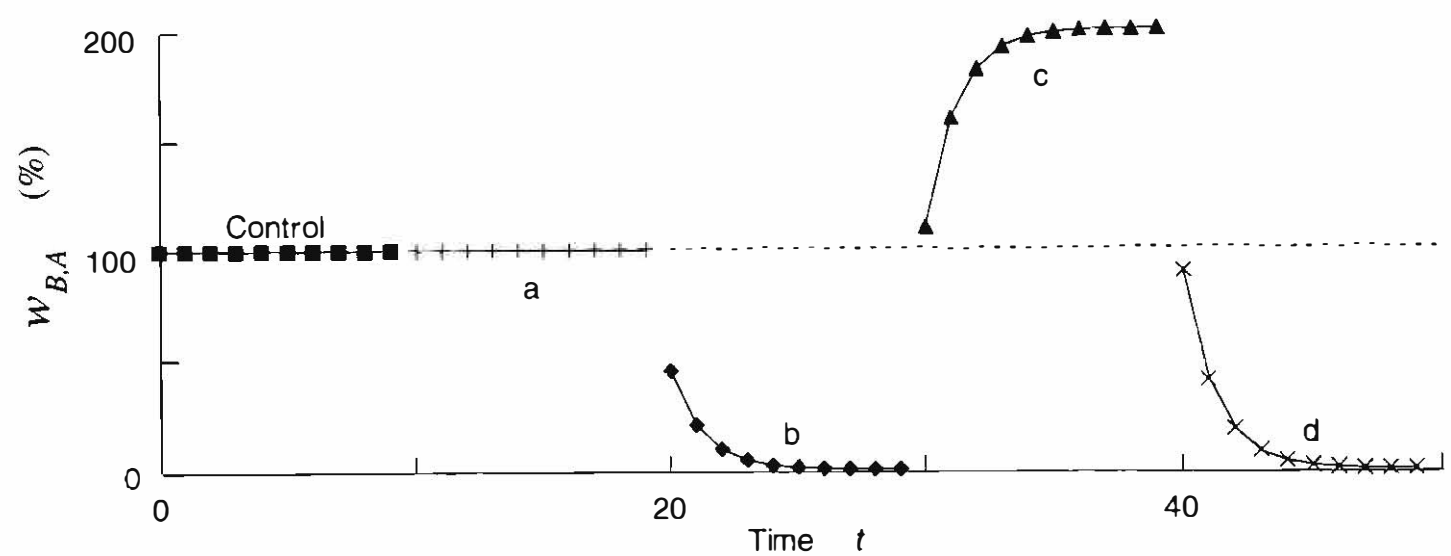
$$w_{B,A}(t+1) = w_{B,A}(t) + [\varepsilon G(y_B(t)) (c y_A(t) - w_{B,A}(t))] \quad (4.14)$$

where G is some measure of the net postsynaptic depolarization or activity and c is a positive constant. (In the simplest case $G = y_B(t)$.) Figure 4.5B illustrates the response of equation 4.14 to four phases of stimulation. If the postsynaptic term, G , is equal to zero (phase a in figure 4.5B), then no change in synaptic weight occurs. Thus, the postsynaptic term is permissive of change. If $G > 0$, then the difference $c y_A - w_{B,A}$ determines whether potentiation or depression of synaptic

A. Equation 4.13



B. Equation 4.14



Changes in Synaptic Weight Predicted by Two Forms of the Hebbian Learning Rule

Figure 4.5. (A) Bilinear Hebbian rule (eq. 4.13) suggested by Brown et al. (1990). The changes in synaptic weight $w_{B,A}$ that are predicted by equation 4.13 are illustrated during four phases of stimulation ($\epsilon = 0.00385$, $\beta = 0.005$, $\gamma = 0.005$, and $\delta = 1$, in all four phases and control value of $w_{B,A} = 1$). During phase a, only the presynaptic cell is active ($y_A = 40$ and $y_B = 0$) and $w_{B,A}$ undergoes homosynaptic depression. During phase b, only the postsynaptic cell is active ($y_A = 0$ and $y_B = 40$) and $w_{B,A}$ again is depressed. During phase c, conjunction activity ($y_A = 40$ and $y_B = 40$) induces potentiation of $w_{B,A}$. During phase d, noncoincident activity ($y_A = 0$ and $y_B = 40$) again induces depression. (B) Hebbian rule incorporating reversible synaptic modifications and asymptotic limits on both potentiation and depression of synaptic weight (eq. 4.14), suggested by Levy (1985). The changes

weight occurs. When $c y_A < w_{B,A}$, depression occurs (phase b). When $c y_A > w_{B,A}$, potentiation occurs (phase c). Finally, the potentiation can be reversed when $c y_A < w_{B,A}$ and $G > 0$ (phase d). This reversal of potentiation is termed depotentiation (Levy 1985; see also Fujii et al. 1991). Learning rules that allow both increases and decreases in synaptic strength avoid eventual saturation and allow synapses to function throughout a broader dynamic range.

An additional feature that can be incorporated into these mathematical expressions of the Hebbian rule is a term describing the temporal specificity. As discussed above, presynaptic activity appears to produce a “trace” that allows the synapses to remain eligible for potentiation for several tens of milliseconds following cessation of presynaptic stimulation. A common method for incorporating a “trace” is by offsetting the presynaptic term by a time τ (for examples and additional discussion see Brown et al. 1990; Klopff 1988, 1989; Sejnowski and Tesauro 1989; Sutton and Barto 1981; Tesauro 1986). A general equation for the Hebbian rule that incorporates the differential effects of several values of τ can be written as follows:

$$\Delta w_{B,A}(t) = \varepsilon \sum_{\tau=0}^k F(y_A(t-\tau)) G(y_B(t)) \quad (4.15)$$

where $F(y_A(t-\tau))$ represents a function of the presynaptic stimulus “trace” and $G(y_B(t))$ represents a function of the postsynaptic activity. Several forms of the function F have been proposed. For example, F can equal the previous level of presynaptic activity, $c_\tau y_A(t-\tau)$ where c_τ is a positive coefficient that determines the contribution of each offset, or F can equal the change in the level of presynaptic activity, $c_\tau (y_A(t) - y_A(t-\tau))$. Similarly, there are a number of possible forms for G , such as the level of postsynaptic activity, $y_B(t)$, or the change in the level of postsynaptic activity. The use of changes in the level of activity is a way of reversing the sign of the synaptic modification (e.g., potentiation vs. depression) and thus avoids the problem of saturation (Klopff 1988, 1989).

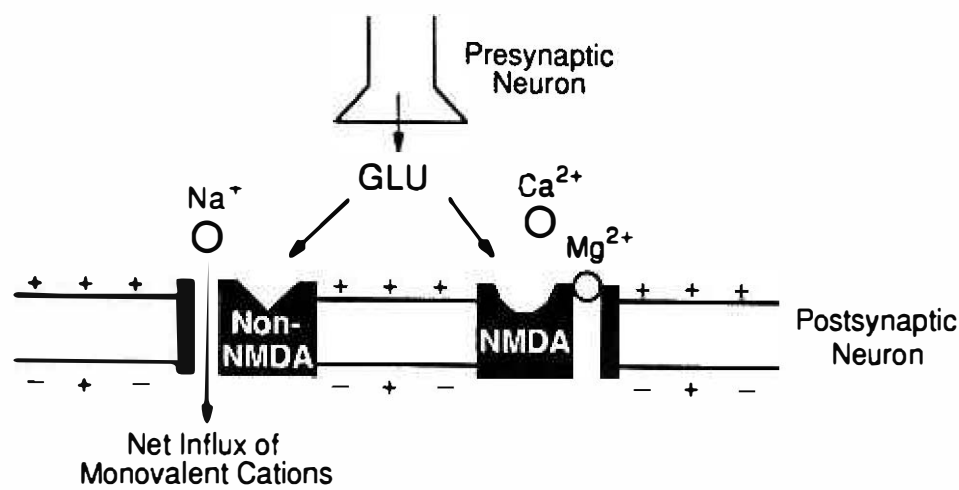
In addition to the mathematical equations for the Hebbian rule discussed above (equations 4.12 through 4.15), a number of more biologically realistic and dynamic models of Hebbian synaptic plasticity have been developed (Gally et al. 1990; Gamble and Koch 1987; Holmes and Levy 1990; Kitajima and Hara 1990, 1991; Lisman 1989; Wickens 1988;

in $w_{B,A}$ that are predicted by equation 4.14 are illustrated during four phases of stimulation ($\varepsilon = 0.011$, and $c = 0.04$ in all four phases; initial value of $w_{B,A} = 1$). During phase a, only the presynaptic cell is active ($y_A = 50$ and $y_B = 0$) and $w_{B,A}$ remains unchanged. During phase b, only the postsynaptic cell is active ($y_A = 0$ and $y_B = 50$) and $w_{B,A}$ is depressed. During phase c, conjunctive activity ($y_A = 50$ and $y_B = 50$) induces potentiation of $w_{B,A}$. During phase d, noncoincident activity ($y_A = 0$ and $y_B = 50$) “depotentiates” $w_{B,A}$.

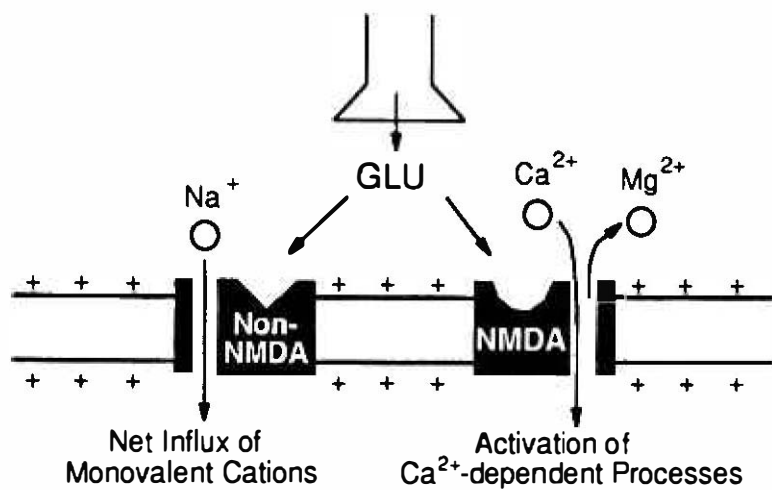
Zador et al. 1990). These models include formal descriptions of the many complex biophysical and/or biochemical processes that may underlie the induction and maintenance of associative LTP. For example, the central tenet of the Hebbian rule is the requirement for conjunctive presynaptic activity and postsynaptic depolarization in order for synaptic modifications to occur. It is generally agreed that this feature of Hebbian potentiation in CA1 area of the hippocampus is derived from an unusual property of the *N*-methyl-D-aspartate (NMDA) subtype of glutamate receptor: its dual regulation by transmitter and membrane voltage (for reviews see Brown et al. 1988, 1990, 1991; Bliss and Collingridge 1993; Collingridge and Bliss 1987; Collingridge and Singer 1990; Gustafsson and Wigström 1988; Madison et al. 1991; Nicoll et al. 1988; Siegelbaum and Kandel 1991; Teyler and DiScenna 1987; see also Cline 1991). As illustrated in figure 4.6, two events must occur before the ionophore of the NMDA receptor will open: The presynaptic transmitter, glutamate, must bind to the receptor and the Mg^{2+} block of the channel must be removed by a strong postsynaptic depolarization. The subsequent influx of Ca^{2+} through the open NMDA ionophore is critical for the induction of synaptic modifications (Collingridge et al. 1983; Lynch et al. 1983; Malenka et al. 1988). The elevation of postsynaptic Ca^{2+} is believed to trigger a complex sequence of enzymatic reactions that alter the properties of the postsynaptic and/or presynaptic cells (for reviews see Madison et al. 1991; Siegelbaum and Kandel 1991; Tsumoto 1990). (Chapter 2 discusses the consequences of this functioning for neuronal modeling, and chapter 3 points out similarities between the subsequent alteration of presynaptic transmitter release and connectionist models.)

Models that include quantitative descriptions of the NMDA receptor, of voltage-gated ionic channels, of the geometric and electrotonic structure of dendritic spines and process, and of the dynamics of the regulation of intracellular Ca^{2+} predict that short, high-frequency stimuli are more effective in elevating the concentration of Ca^{2+} than longer, low-frequency stimuli (Gamble and Koch 1987; Holmes and Levy 1990; Zador et al. 1990). Moreover, a model of Ca^{2+} -dependent enzymes in the postsynaptic spines predicts that the elevation of Ca^{2+} can trigger either potentiation or depression of synaptic efficacy (Lisman 1989). The "sign" of the modification is determined by the levels of Ca^{2+} : Lower levels of Ca^{2+} are predicted to induce synaptic depression and higher levels are predicted to induce potentiation. Taken together, these models appear to accurately predict the experimental observations that tonic, low-frequency stimulation tends to induce depression or depotentiation, whereas patterned, high-frequency stimulation tends to induce potentiation (Dudek and Bear 1992; Dunwiddie and Lynch 1978; Grover and Teyler 1990; Larson and Lynch 1986; Mulkey and Malenka 1992; Pavlides et al. 1988; Staubli and Lynch 1987, 1990). Moreover, these models provide insights into the spatial and temporal interactions among synapses in the complex dendritic structure of the postsynaptic cell. For

A. Presynaptic Activity in Conjunction with a Small Postsynaptic Depolarization



B. Presynaptic Activity in Conjunction with a Large Postsynaptic Depolarization



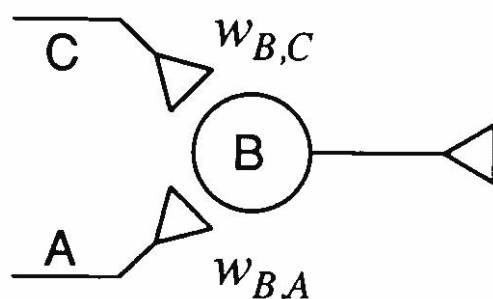
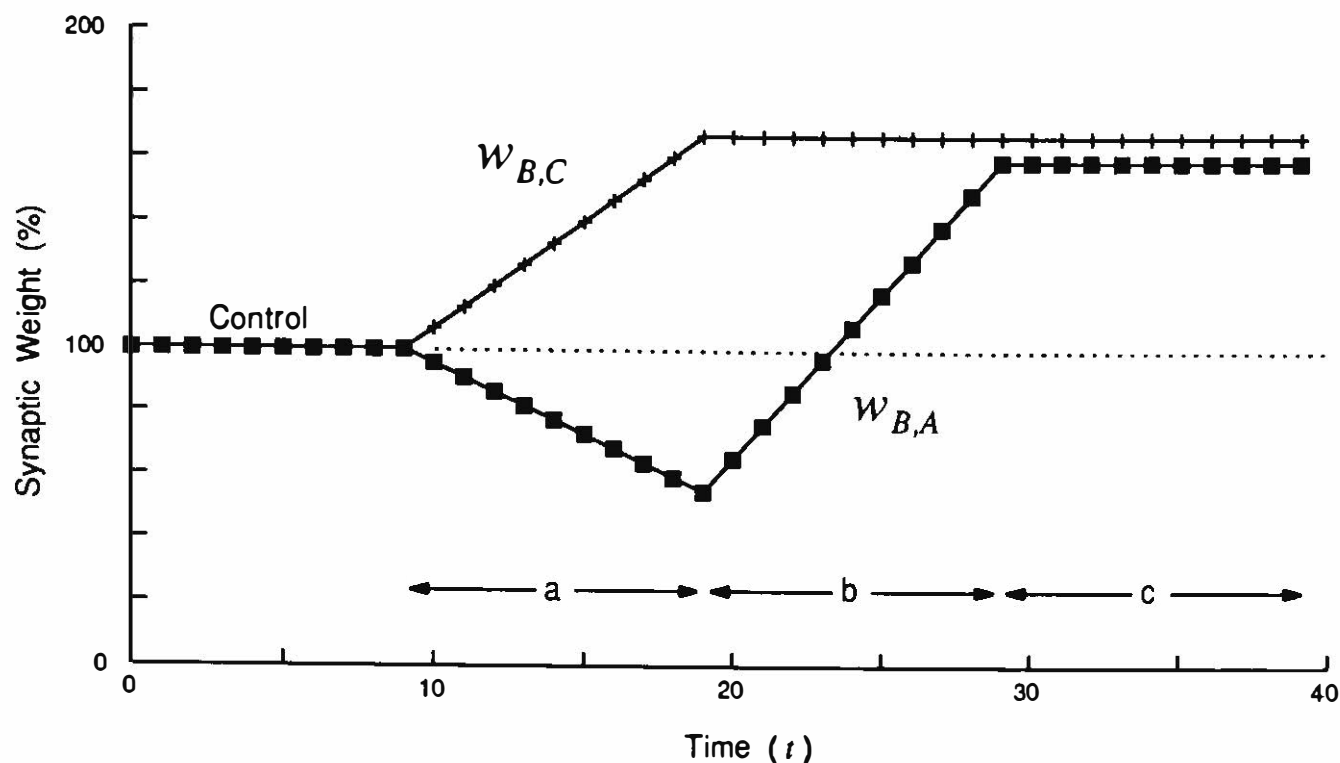
The NMDA Receptor: A Molecular Basis for the Hebbian Learning Rule

Figure 4.6 The dual regulation of the NMDA receptor by the binding of glutamate and by membrane voltage allows for the detection of conjunctive presynaptic and postsynaptic activity. (A) Presynaptic activity alone releases the excitatory transmitter glutamate (GLU), which binds to both NMDA and non-NMDA receptors. A net influx of monovalent cations through ionophores associated with the non-NMDA receptors produces postsynaptic depolarization. A voltage-dependent block by Mg²⁺ of the ionophore associated with the NMDA receptor prevents the postsynaptic influx of Ca²⁺ that is necessary for the induction of synaptic potentiation. (B) Presynaptic release of glutamate in conjunction with a large postsynaptic depolarization removes the Mg²⁺ block of the NMDA channel and allows Ca²⁺ influx into the postsynaptic neuron, activating Ca²⁺-dependent processes that induce synaptic potentiation.

example, recent models indicate that postsynaptic spines play an important role in amplifying and compartmentalizing the Ca^{2+} signals that are crucial for synaptic modification (Gamble and Koch 1987; Holmes and Levy 1990; Koch et al. 1992; Zador et al. 1990).

Anti-Hebbian Rule for Synaptic Modification

An anti-Hebbian rule is one in which synaptic weights are decreased as a result of concurrent presynaptic and postsynaptic activity. As noted by Brown et al. (1990), the term anti-Hebbian is not used consistently in the literature. In addition, the terms “inverse-Hebbian” (Artola et al. 1990) and “reverse-Hebbian” (Donegan et al. 1989) have been used recently to describe synaptic depression resulting from concurrent presynaptic and postsynaptic activity. One form of synaptic plasticity that is generally consistent with the anti-Hebbian rule is LTD at the synapses from parallel fibers to Purkinje cells in the cerebellar cortex (for recent reviews of cerebellar LTD, see Ito 1989a, b, 1991, 1992; Siegelbaum and Kandel 1991). A similar form of anti-Hebbian synaptic depression recently has been observed in regions of the neocortex (Bindman et al. 1988; Hirsch and Crepel 1990). Stimulation of parallel fibers in conjunction with strong postsynaptic depolarizations, which elicit dendritic Ca^{2+} spikes, induces LTD at the stimulated synapses (Crepel and Jaillard 1991; Hirano 1990b; see also Crepel and Krupa 1988; Linden and Connor 1991; Linden et al. 1991). The induction of LTD is believed to require the concurrent increase in postsynaptic levels of Ca^{2+} (Sakurai 1990; see also Hirsh and Crepel 1992) and the activation of glutamate receptors (Kano and Kato 1987; Linden et al. 1991), which mediate parallel fiber to Purkinje cell transmission (Hirano 1990c; Hirano and Hagiwara 1988; Konnerth et al. 1990). These two events initiate a sequence of enzymatic reactions (for reviews see Siegelbaum and Kandel 1991; Ito 1991, 1992), the end result of which is to desensitize those glutamate receptors activated by the parallel fiber stimulation, which in turn depresses the efficacy of parallel fiber transmission. In addition to LTD, LTP of parallel fibers can be induced by either stimulating the parallel fibers alone, or stimulating them in conjunction with postsynaptic hyperpolarizations or with weak depolarizations, which elicit Na^+ but not Ca^{2+} spikes (Crepel and Jaillard 1991; Hirano 1990a, b; Sakurai 1987). Thus, parallel fiber synapses can undergo either potentiation or depression and the determining factor is believed to be the level of Ca^{2+} in the postsynaptic cells: High levels of Ca^{2+} induce LTD, whereas low levels of Ca^{2+} allow for LTP. Interestingly, this is in contrast to plasticity in the CA1 region of the hippocampus, where high levels of postsynaptic Ca^{2+} are necessary for LTP and low levels are believed to induce LTD. There have been various attempts to model aspects of the cerebellar circuitry and to include mathematical expressions for plasticity at parallel fiber to Purkinje cell synapses (Chapeau-Blondeau and Chauvet 1991; Chauvet 1986;



- a. Cells C and B positively correlated
Cells A and B negatively correlated
- b. Cells C and B uncorrelated
Cells A and B positively correlated
- c. Cells A, C and B Uncorrelated

Changes in Synaptic Weight Predicted by the Covariance Learning Rule

Figure 4.7 In the circuit shown, postsynaptic neuron B is innervated by two presynaptic neurons, A and C. The changes in synaptic weights predicted by the covariance rule, equation 4.16 (Sejnowski 1977), are illustrated during three phases of stimulation. During phase a, activity in cell A is negatively correlated with postsynaptic activity and its synaptic weight $w_{B,A}$ decreases ($y_A = 17$ and $y_B = 25$), whereas activity in cell C is positively correlated with the postsynaptic cell ($y_C = 25$ and $y_B = 25$), and its synaptic weight $w_{B,C}$ potentiates. During phase b, the activity in cell A is positively correlated with the postsynaptic cell, and its synaptic weight potentiates ($y_A = 27$ and $y_B = 25$), whereas activity of cell C is uncorrelated with the postsynaptic cell, and there is no change in its synaptic weight ($y_C = 20$ and $y_B = 25$). During phase c, the activities in both presynaptic cells are uncorrelated with activity in the postsynaptic cell, and there is no change in either of their synaptic weights ($y_A = 20$, $y_C = 20$, and $y_B = 20$). In all three phases, $\epsilon = 0.003$, $\bar{y}_A = 20$, $\bar{y}_B = 20$, and $\bar{y}_C = 20$; initial values of both $w_{B,A}$ and $w_{B,C} = 1$.

Desmond and Moore 1988; Houk et al. 1990; Moore and Blazis 1989; Moore et al. 1989). Although the role that the cerebellum might play in motor learning is unclear (Bloedel et al. 1991; Donegan et al. 1989; Greenough and Anderson 1991; Thompson 1986, 1989; Wagner and Donegan 1989; Yeo 1991), these theoretical studies predict that the cerebellar circuitry and its related synaptic plasticity, in principle, can perform complex information processing and contribute to motor learning.

Covariance Rule for Synaptic Modification

Stent (1973) offered an extension to the Hebbian rule:

When the presynaptic axon of cell A repeatedly and persistently fails to excite the postsynaptic cell B while cell B is firing under the influence of other presynaptic axons, metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is decreased.

This rule further specifies that synaptic depression results from "an asynchronous activity pattern of the two cells"; not simply the disuse of the presynaptic axon. Aspects of this generalization of the Hebbian rule can be found in a learning rule that is referred to as the "covariance rule" (Sejnowski 1977; Sejnowski and Stanton 1990; see also Chauvet 1986; Cooper 1986; Dayan and Willshaw 1991; Reilly and Cooper 1990; Rubner and Schulten 1990). According to the covariance rule, changes in synaptic weight are proportional to the covariance between the firing rates of the presynaptic and postsynaptic cells:

$$w_{B,A}(t+1) = w_{B,A}(t) + [\epsilon (y_A(t) - \bar{y}_A) (y_B(t) - \bar{y}_B)] \quad (4.16)$$

where $\bar{y}_A$ and $\bar{y}_B$ as before measure the sustained levels of activity in the presynaptic and postsynaptic cells, respectively. Thus, synaptic weight should increase if the pre- and postsynaptic activities are positively correlated, decrease if they are negatively correlated, and remain unchanged if they are uncorrelated. Figure 4.7 illustrates a numerical example showing how the synaptic weights of two presynaptic cells, A and C, can be governed by the covariance rule during three phases of stimulation. During the first phase (phase a in figure 4.7), activity in cell A is negatively correlated with postsynaptic activity and its synaptic weight ($w_{B,A}$) decreases. In contrast, the activity in cell C is positively correlated with the postsynaptic cell, and its synaptic weight ($w_{B,C}$) potentiates. During the second phase (phase b), the activity in cell A is positively correlated with the postsynaptic cell, and its synaptic weight potentiates. During this phase, activity of cell C is uncorrelated with the postsynaptic cell, and there is no change in its synaptic weight. Finally, in phase c, the activities in both presynaptic cells are uncorrelated with activity in the postsynaptic cell, and there is no change in either of their synaptic weights. Thus, the covariance rule overcomes a limitation inherent with nondecremental learning rules such as equation 4.12 and allows the weight of a synapse to be used throughout its dynamic range. In

addition, the covariance rule solves the problem of automatic growth in synaptic weight resulting from random coincidence of presynaptic and postsynaptic activity.

Recently, there have been reports of a form of synaptic depression in the hippocampus that appears to be a good candidate for the covariance rule (Sejnowski and Stanton 1990; Stanton and Sejnowski 1989; Stanton et al. 1989, 1991; see, however, Goldman et al. 1990). For example, a neural analog of the covariance rule was examined by stimulating two distinct inputs to CA1 pyramidal neurons, a strong input and a weak input. Stimulation of the strong input induced potentiation in the stimulated synapses without significantly affecting the unstimulated weak input. Stimulation of the weak input alone did not induce any changes in the synaptic efficacy of either input. If the weak input was stimulated out of phase with the strong input, however, depression was induced in the synapses stimulated by the weak input. This phenomenon was referred to as associative LTD (Stanton and Sejnowski 1989). Recently, a similar form of synaptic depression induced by asynchronous presynaptic and postsynaptic activity has been reported at vertebrate neuromuscular synapses in culture (Dan and Poo 1992). If the weak and strong inputs were stimulated in phase (simultaneously), potentiation was induced in the synapses of the weak input. The voltage of the postsynaptic cell may be an important determinant of this form of synaptic depression (Sejnowski and Stanton 1990). If stimulation of weak input was paired with intracellular depolarization of the postsynaptic cell, potentiation was induced in the synapses of the weak input. This is the Hebbian rule for synaptic potentiation. If stimulation of weak input was paired with hyperpolarization of the postsynaptic cell, however, depression was induced. A similar form of synaptic depression that results from pairing presynaptic activity with postsynaptic hyperpolarization recently has been observed in the dentate gyrus area of the hippocampus (Xie et al. 1992) and the visual cortex (Friedlander et al. 1993). Thus, there is evidence of a form of synaptic plasticity that is governed by the covariance rule, in that identical presynaptic stimulations of the weak input can produce either potentiation or depression depending on the phase relationship between presynaptic and postsynaptic activity.

Bidirectional/Variable Threshold Rule for Synaptic Modification

Bienenstock et al. (1982; see also Bear et al. 1987; Clothiaux et al. 1991; Cooper 1986; Cooper et al. 1979, 1985, 1990; Intrator and Cooper 1992) proposed a rule for synaptic modification that, in theory, might underlie the self-organization and learning in neural networks. This modification rule, sometimes referred to as the Bienenstock, Cooper, and Munro (BCM) rule, is predicated on the notion that individual synapses are capable of bidirectional modifications (both depression and potentiation) and whether the synaptic strength increases or decreases depends

upon the magnitude of the postsynaptic response as compared to a variable threshold θ . A general mathematical expression for this rule is:

$$\Delta w_{B,A}(t) = \phi(y_B(t), \theta(t)) y_A(t), \quad (4.17)$$

where ϕ is a scalar function of the postsynaptic activity that changes sign at a threshold θ and where y_B and y_A again represent activity in the postsynaptic and presynaptic cells, respectively. The precise form of ϕ is not critical as long as it has certain general characteristics:

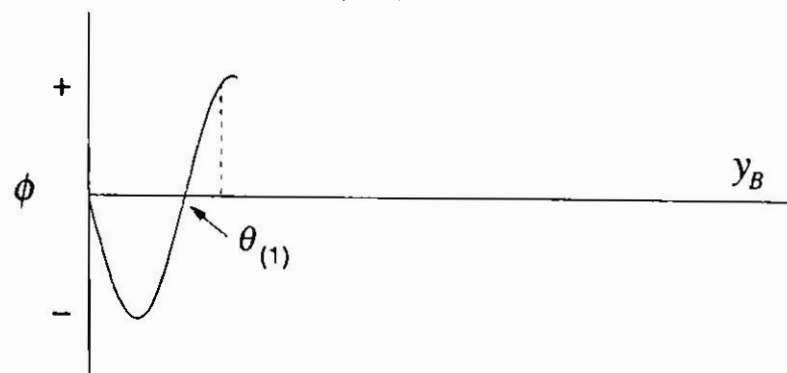
$$\phi = 0 \text{ for } y_B = 0; \quad \phi < 0 \text{ for } y_B < \theta; \quad (4.18a)$$

$$\phi = 0 \text{ for } y_B = \theta; \quad \phi > 0 \text{ for } y_B > \theta. \quad (4.18b)$$

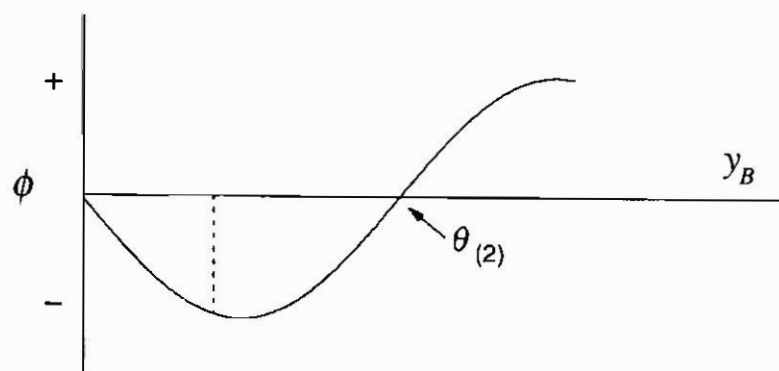
An arbitrary function that satisfies these general features is illustrated in figure 4.8 for two different values of the modification threshold θ . The two essential features of this rule are: first, ϕ changes sign at θ and second, the value of θ is not fixed. Thus, this rule predicts that a given presynaptic conditioning stimulus can produce such categorically different responses as potentiation, no change or depression depending on the magnitude of the postsynaptic response that the input elicits and on the current value of θ . For example, if the current threshold for synaptic modification is low (figure 4.8A, $\theta_{(1)}$), then a given presynaptic input that elicits the postsynaptic response indicated by the dotted line in figure 4.8A would induce synaptic potentiation; ϕ is positive. If the current threshold for synaptic modification is high (figure 4.8B, $\theta_{(2)}$), however, then the same presynaptic input and the same postsynaptic response (dotted line in figure 4.8B) would induce synaptic depression; ϕ is negative.

Recent experiments that examined synaptic plasticity in slices of visual cortex have provided support for some of the essential features of this rule. Artola et al. (1990; see also Reiter and Stryker 1988; Singer et al. 1990) found that identical presynaptic conditioning stimuli can lead to either no change, depression or potentiation of the activated synapses and that the critical determinant of the nature of synaptic modification appeared to be the level of postsynaptic depolarization or activity. Thus, presynaptic activity that was paired with postsynaptic hyperpolarization led to no change of synaptic efficacy, whereas presynaptic activity that was paired with moderate postsynaptic depolarizations led to synaptic depression, and presynaptic activity that was paired with stronger postsynaptic depolarizations led to synaptic potentiation. The molecular mechanism for the modification threshold θ is believed to involve the level of Ca^{2+} in the postsynaptic cell (Cooper et al. 1990; Kimura et al. 1990; Singer 1983, 1987, 1990). Moderate postsynaptic depolarizations are believed to activate voltage-dependent Ca^{2+} channels resulting in moderate increases in postsynaptic Ca^{2+} levels, which in turn lead to synaptic depression (see Bröcher et al. 1992). Stronger postsynaptic depolarizations are believed to activate Ca^{2+} influx through both voltage-dependent Ca^{2+} channels and NMDA receptors, which results in

A. Low Threshold for Synaptic Modification



B. High Threshold for Synaptic Modification



The Modification Function $\phi(y_B, \theta)$ Shown for Two Values of θ

Figure 4.8 ϕ is a theoretical function relating synaptic weight change to postsynaptic activity y_B . The sign of ϕ changes at θ , such that ϕ is negative (synaptic depression) for $y_B < \theta$, and ϕ is positive (synaptic potentiation) for $y_B > \theta$. The value of θ is variable. (A) Plot of ϕ for a low modification threshold, $\theta_{(1)}$. (B) Plot of ϕ for a high modification threshold, $\theta_{(2)}$. Thus, a presynaptic conditioning stimulus that produces the same level of postsynaptic activity (dotted lines) would produce synaptic potentiation if θ has a low value ($\theta_{(1)}$), but would produce synaptic depression if θ has a high value, ($\theta_{(2)}$). (Modified from Bienenstock et al. 1982, and from Yang and Faber 1991.)

greater levels of postsynaptic Ca^{2+} and synaptic potentiation (see Artola and Singer 1990; Artola et al. 1990; Bear et al. 1990; Singer et al. 1990).

In the original formulation of this learning rule (equation 4.17), it was proposed that θ vary as a nonlinear function of the average output ($\bar{y}_B$) of the postsynaptic cell (Bienenstock et al. 1982; see also Clothiaux et al. 1991):

$$\theta = F(\bar{y}_B) = (\bar{y}_B)^2. \quad (4.19)$$

It should be noted that Bienenstock et al. (1982; see also Singer 1983, 1987) proposed that θ may also be influenced by global signals, such as modulatory systems; i.e., $\theta = F(\bar{y}_B, y_M)$. Recent experiments have illustrated that modulatory transmitters can have a permissive role for the induction of synaptic plasticity in the visual cortex (Greuel et al. 1988; Singer 1990). Recently Yang and Faber (1991) suggested as an alternative to equation 4.19 that the modification threshold θ might vary as a function of the initial level of synaptic efficacy: $\theta = F(w_{B,A})$. For example, potentiation of the synaptic inputs to the Mauthner cell of the goldfish can be induced by conjunctive presynaptic and postsynaptic activity

(Yang et al. 1990). If the synaptic efficacy is first elevated, however, the conditioning stimulus, which usually induces potentiation, produces synaptic depression instead (Yang and Faber 1991). Thus, the sign of the synaptic modification was changed by varying the level of synaptic efficacy. These results suggest the hypothesis that the modification threshold θ is increased for synapses with an initial efficacy that is elevated. Additional support for the hypothesis that θ may vary as a function of prior synaptic activity recently has been observed in the hippocampus (Huang et al. 1992).

The nature of the function that controls the modification threshold can have important implications for information processing and storage within single cells. Since $\bar{y}_B$ reflects postsynaptic activity, the value of θ would be the same for all presynaptic neurons that converge onto the cell. In contrast, if θ were a function of presynaptic efficacy, then each synapse could have a different value of θ . For example, consider a postsynaptic neuron, B , with one strong synaptic input, $w_{B,C}$, and one weak synaptic input, $w_{B,A}$. If θ were a function of $\bar{y}_B$, then the stronger synapse could maintain $\bar{y}_B$, hence θ , at a high level for both synapses, and thus could prevent the weaker synapse from potentiating. In contrast, if θ were a function of synaptic efficacy, then the value of θ should be less for the weaker synapse and greater for the stronger synapse.

Activity-Dependent Neuromodulation of Synaptic Efficacy

Activity-dependent neuromodulation is a form of associative synaptic plasticity in which presynaptic activity in conjunction with the presence of a modulatory transmitter produces a pairing-specific potentiation (or depression) of synaptic efficacy. This represents an extension of rule 4 (see figure 4.1) by requiring conjunctive activity in the two pathways. Thus, activity-dependent neuromodulation can be represented by the general expression:

$$\Delta w_{B,A}(t) = F(y_A, y_M), \quad (4.20)$$

which relates changes in the efficacy of the synapse from neuron A to neuron B to activity in the presynaptic neuron y_A and in a modulatory neuron y_M . The key feature of this synaptic modification rule is that the conjunction of presynaptic activity and the presence of the modulatory transmitter produces a significantly greater change in synaptic efficacy than a simple summation of the effects that presynaptic activity alone (rule 1) and the modulatory transmitter alone (rule 4) have on synaptic efficacy. Activity-dependent neuromodulation has been observed in a number of excitatory synapses, including projections from the entorhinal cortex (perforant path) to granule cells in the dentate gyrus of the hippocampus (Williams et al. 1989), the mossy fiber synapses that arise from the dentate gyrus and terminate on the CA3 pyramidal neurons in the hippocampus (Hopkins and Johnston 1984, 1988; Johnston et al.

1989), Sch/comm projections to CA1 pyramidal neurons in the hippocampus (O'Dell et al. 1991; Small et al. 1992), crayfish neuromuscular synapses (Breen and Atwood 1983), and sensorimotor synapses of *Aplysia* (Buonomano and Byrne 1990; Hawkins et al. 1983; Small et al. 1989; Walters and Byrne 1983).

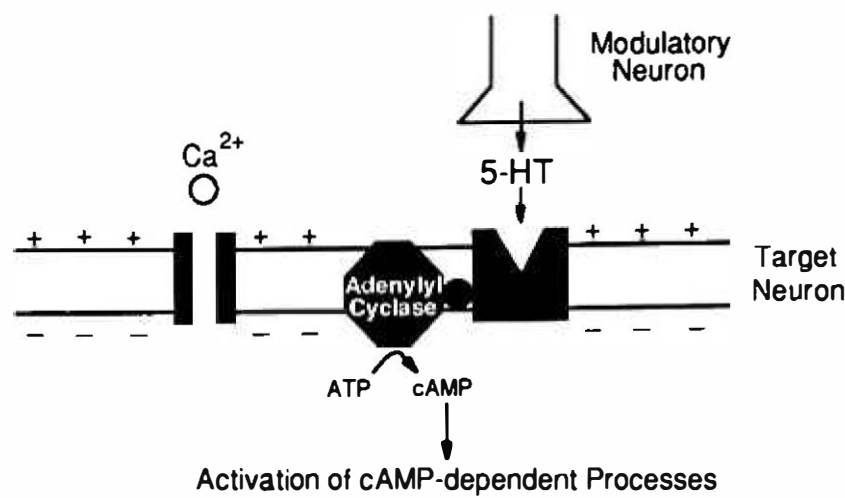
Gluck and Thompson (1987; see also Donegan et al. 1989) suggested a mathematical formalism to describe the activity-dependent neuromodulation of sensorimotor synapses in *Aplysia*:

$$\Delta w_{B,A}(t) = \varepsilon y_M(t) F(y_A(t - \tau)) (\lambda_{\max} - w_{B,A}(t)) - (\gamma y_A(t)) \quad (4.21)$$

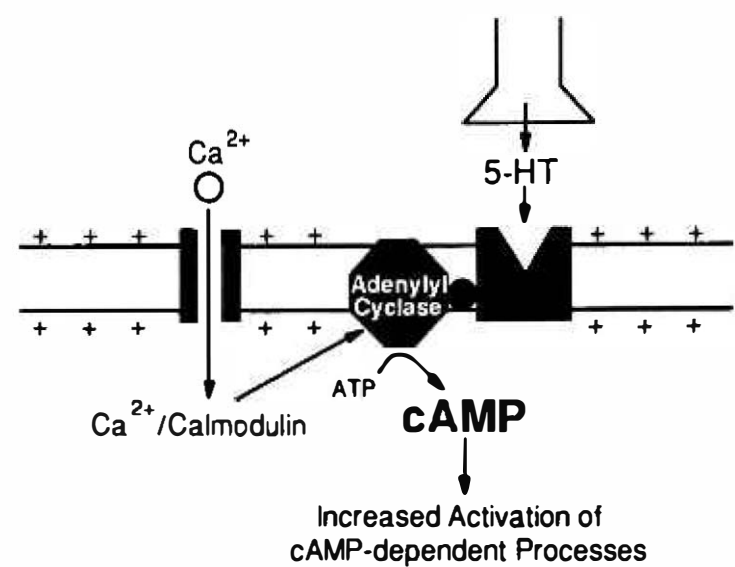
where y_M is a measure of activity in a modulatory neuron, $F(y_A(t - \tau))$ is a function describing a stimulus “trace” that is produced by presynaptic activity, and $(-\gamma y_A(t))$ represents a form of nonassociative homosynaptic depression that is induced by presynaptic activity. The temporal specificity of the associative interaction between the activities in the presynaptic neuron, y_A , and the modulatory neuron, y_M , is governed by the function $F(y_A(t - \tau))$. The suggested form of F was the product of exponentially decaying and rising functions. Thus, immediately after activation of the presynaptic neuron, the trace function F is 0, but then rises quickly to a peak and then slowly decreases. Equation 4.21 predicts that presynaptic activity ($y_A > 0$) in the absence of the modulator ($y_M = 0$) would result in nonassociative homosynaptic depression, and that activity in the modulatory neuron ($y_M > 0$) in the absence of presynaptic activity ($y_A = 0$) would not alter synaptic weight. Moreover, the temporal specificity function F predicts that simultaneous activity in the modulatory and presynaptic neurons would not induce synaptic potentiation. Rather, activity in the modulatory neuron would be most effective at inducing potentiation a short time after activity in the presynaptic neuron, when F reaches its peak.

Although equation 4.21 captures some features of activity-dependent neuromodulation, this formalism is an oversimplification of the cellular mechanisms contributing to this form of associative plasticity at sensorimotor synapses in *Aplysia*. Many of the biophysical and biochemical processes that contribute to activity-dependent neuromodulation in *Aplysia* have been determined (for reviews, see Abrams and Kandel 1988; Byrne 1987; Byrne and Crow 1991; Byrne et al. 1991a). For example, a key feature of this associative rule for synaptic plasticity is that the conjunction of presynaptic activity and the presence of the modulatory transmitter produces a significantly greater potentiation of synaptic efficacy than either treatment alone. It is generally believed that this feature of activity-dependent neuromodulation is derived from the dual regulation of the enzyme adenylyl cyclase in the sensory neurons by the modulatory transmitter serotonin and by intracellular Ca^{2+} /calmodulin (Abrams and Kandel, 1988; Abrams et al. 1991; Eliot et al. 1989; Ocorr et al. 1985; Yovell and Abrams 1992; see also Dudai 1987). As illustrated in figure 4.9, binding of serotonin to its receptor stimulates adenylyl

A. Modulatory Transmitter Alone



B. Modulatory Transmitter in Conjunction with Depolarization



Adenylyl Cyclase: A Molecular Basis for the Activity-Dependent Neuromodulation Learning Rule

Figure 4.9 The enzyme adenylyl cyclase detects the conjunctive activity in a neuron and the modulatory effects produced in that neuron by a modulatory pathway. (A) Activity in the modulatory neuron releases the transmitter serotonin (5-HT). Binding of serotonin to its receptor stimulates adenylyl cyclase and elevates levels of cAMP in the target neuron. The increased levels of cAMP, in turn, activate cAMP-dependent processes that result in nonassociative heterosynaptic potentiation. (B) An influx of Ca^{2+} during action potentials in the target neuron interacts with a Ca^{2+} /calmodulin-sensitive component of adenylyl cyclase and significantly amplifies the production of cAMP elicited by the modulatory transmitter, and hence the activation of cAMP-dependent processes.

cyclase and elevates levels of cAMP in the sensory neurons (figure 4.9A). The increased levels of cAMP, in turn, activate cAMP-dependent processes, such as cAMP-dependent protein kinase A, that modulate a number of properties of the sensory neurons, including ion channels and secretory machinery, that contribute to synaptic potentiation. Thus, the release of the modulatory transmitter alone induces nonassociative heterosynaptic potentiation (for reviews, see Byrne et al. 1991c; Kandel and Schwartz 1982). The influx of Ca^{2+} during presynaptic activity (figure 4.9B) interacts with a Ca^{2+} /calmodulin-sensitive component of adenylyl cyclase and amplifies the production of cAMP elicited by the modulatory transmitter. This dual activation of adenylyl cyclase requires that the Ca^{2+} and modulatory transmitter temporally overlap. Thus, enhanced synthesis of cAMP by conjunctive presynaptic activity and the modulatory transmitter induces a pairing-specific potentiation of synaptic efficacy.

Two additional biologically realistic and dynamic models of activity-dependent neuromodulation have been developed that reflect aspects of the cellular mechanisms underlying this form of associative synaptic plasticity in sensory neurons of *Aplysia*.

Hawkins (1989a, b) developed a relatively simple model of seven differential equations that describe three phenomena: (1) nonassociative homosynaptic depression solely as a function of Ca^{2+} channel inactivation; (2) nonassociative heterosynaptic potentiation as a function of cAMP-dependent increases in spike duration, and therefore Ca^{2+} current; and (3) activity-dependent neuromodulation as a function of Ca^{2+} -dependent enhancement of cAMP synthesis.

Gingrich and Byrne (1984, 1985, 1987) developed a somewhat more detailed model. Some of the general features of the Gingrich and Byrne model are illustrated in figure 4.10. The details of this model have been described previously (Buonomano et al. 1990; Byrne and Gingrich 1989; Byrne et al. 1989; Gingrich and Byrne 1985, 1987; Gingrich et al. 1988; Raymond et al. 1992). The model contains differential equations describing two pools of transmitter, a releasable pool and a storage pool. During a presynaptic action potential, an influx of Ca^{2+} through voltage-dependent Ca^{2+} channels causes the release of transmitter. Thus, the amount of transmitter that is released is a function of both the dynamics of Ca^{2+} influx and the number of vesicles in the releasable pool. As a consequence of release, the releasable pool is depleted. In order to offset depletion, transmitter is delivered (mobilized) from the storage pool to the releasable pool. The mobilization process is regulated, in part, by the intracellular concentrations of cAMP and Ca^{2+} . The influx of Ca^{2+} during simulated action potentials leads to the release of transmitter and accumulation of intracellular Ca^{2+} . The pool of intracellular Ca^{2+} is divided into a submembrane compartment and an interior compartment. The Ca^{2+} within the submembrane compartment regulates the release of transmitter, whereas the Ca^{2+} within the interior compartment contrib-

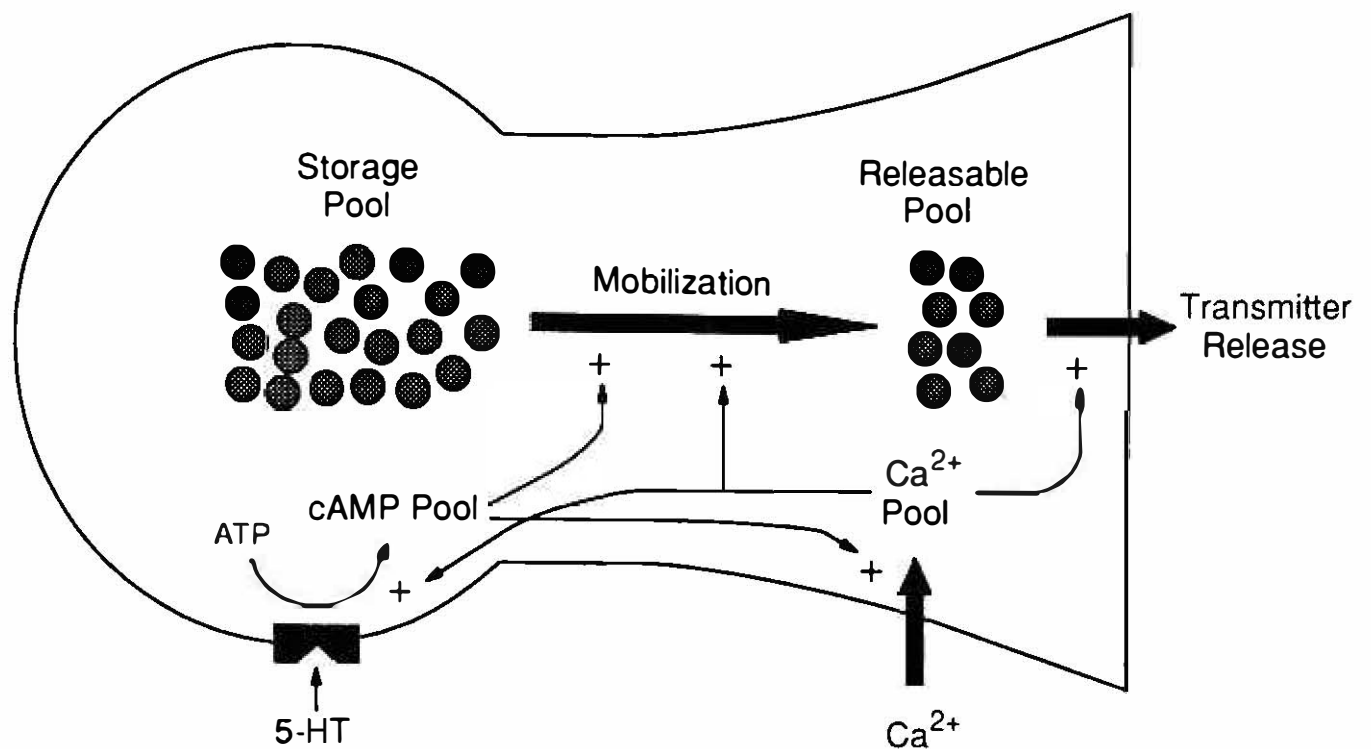
utes to the regulation of mobilization and to the regulation of cAMP synthesis (see below). Two fluxes remove Ca^{2+} from the interior compartment: one that represents active buffering of Ca^{2+} by organelles and one that represents diffusion of Ca^{2+} into an innermost compartment that serves as a Ca^{2+} sink.

The Gingrich and Byrne model also includes differential equations describing the concentration of cAMP and its effects on the release of transmitter. The modulatory transmitter, 5-HT, activates adenylyl cyclase, which leads to synthesis of cAMP. Increased levels of cAMP contribute to an increase in the duration of the action potential, which results in an increased influx of Ca^{2+} during a subsequent action potential in the sensory neuron, and hence, an increase in the release of transmitter from the sensory neuron. Empirical results indicate that the activity of adenylyl cyclase is also regulated by intracellular levels of Ca^{2+} (see above). In the model, an influx of Ca^{2+} during presynaptic spikes that precede the modulatory transmitter, primes the cyclase and amplifies the subsequent stimulation of cAMP synthesis, which in turn, leads to a pairing-specific enhancement of synaptic efficacy.

This dynamic model accurately simulates many aspects of empirically observed synaptic plasticity in sensory neurons of *Aplysia*. For example, this model simulates nonassociative homosynaptic depression and serotonin-induced heterosynaptic potentiation. In the model, this homosynaptic depression is due to the combined effects of Ca^{2+} -current inactivation and depletion of the releasable pool. Heterosynaptic potentiation is due to the combined effects of transmitter mobilization and cAMP-mediated broadening of the presynaptic action potential. In addition, this model simulates features of activity-dependent neuromodulation, including pairing-specific enhancement of synaptic potentiation and the requirement for a close temporal association between presynaptic activity and the presence of modulatory transmitter. In the model, neuromodulation is a function of the interstimulus interval (ISI) between presynaptic activity and application of the modulatory transmitter. Activity-dependent neuromodulation is optimal for an ISI of about 200 ms, in which presynaptic activity begins 200 ms before the application of the modulatory transmitter, whereas longer ISIs are less effective. This ISI-dependence of the model is a direct consequence of the kinetics of the buffering of intracellular Ca^{2+} . Thus, the elevation of intracellular Ca^{2+} produced by the presynaptic activity serves as a stimulus “trace” that becomes associated with the closely paired application of the modulatory transmitter.

CONCLUSIONS

Several general themes have been highlighted in this chapter. First, many diverse forms of activity-dependent synaptic plasticity exist. Indeed, individual neurons often express more than one form. Second, the in-



Gingrich and Byrne Model for Activity-Dependent Neuromodulation

Figure 4.10 The model contains equations describing two pools of transmitter, a releasable pool and a storage pool. During a simulated action potential, an influx of Ca^{2+} through voltage-dependent Ca^{2+} channels causes the release of transmitter. Transmitter is delivered (mobilized) from a storage pool to the releasable pool. The mobilization process is regulated, in part, by the intracellular concentrations of cAMP and Ca^{2+} . The modulatory transmitter 5-HT activates adenylyl cyclase, which leads to increased synthesis of cAMP. Increased levels of cAMP contribute to an increase in the duration of the action potential, and hence, an increased influx of Ca^{2+} . In the model, an influx of Ca^{2+} during presynaptic spike activity that precedes the modulatory transmitter, primes the cyclase and amplifies the subsequent stimulation of cAMP synthesis. Thus, convergence at adenylyl cyclase of the Ca^{2+} signal and the modulatory transmitter results in an associative amplification of the synthesis of cAMP, which in turn leads to a pairing-specific enhancement of synaptic potentiation. (Modified from Gingrich and Byrne 1987.)

duction of synaptic plasticity is governed by a complex set of spatial and temporal rules that can be imposed by the presynaptic neuron, postsynaptic neuron, neighboring synapses, or by modulatory neurons. Third, in neurons where multiple forms of synaptic plasticity are expressed, these rules can allow for the selective induction of a particular form of synaptic modification or for different regions of a single cell to act independently in the processing and storage of information. Moreover, these rules can allow for the dissociation of information storage from information processing within a single neuron.

Finally, two approaches have been used to describe and model synaptic plasticity quantitatively: biologically realistic, dynamic models and relatively simple, abstract mathematical equations. Each approach has its advantages and disadvantages. The relatively simple equations have the advantage of being computationally efficient; thus they can be readily incorporated into large-scale simulations of neural networks. In addition, they capture or provide a caricature of many of the more general features of the plasticity. These equations, however, do not

incorporate some features of synaptic plasticity that are likely to be important in neuronal information processing and storage, such the stimulus specificity for the induction and reversal of synaptic plasticity or complex spatial and temporal synaptic interactions (nor can they be readily modified to do so). The more biologically realistic models can more accurately predict how different forms of plasticity might emerge and interact within a single synapse, how these different forms of synaptic plasticity can be selectively induced by specific patterns of stimulation, the spatial and temporal interactions among synapses that converge onto overlapping dendritic regions, and the striking nonlinearities that characterize synaptic information processing and storage. Unfortunately, these biologically realistic models tend to be computationally intensive and thus are slow and cumbersome when used for large-scale simulations of neural networks.

Given the diversity of learning rules described in this chapter, an obvious question is, What types of learning do they mediate? Are some types of rules used selectively for some forms of learning (e.g., motor learning and skills) and other rules used for more cognitive other forms of learning (for discussion, see Mishkin et al. 1984; Squire 1987)? Answers to these questions are not yet available, but it seems likely that the consequences of a learning rule will depend very much on the circuit into which it is embedded. Indeed, any one learning rule may mediate different forms of learning. Several qualitative and quantitative modeling studies have demonstrated this principle (Baxter et al. 1991; Brindley 1967; Buonomano et al. 1990; Burke 1966; Byrne et al. 1991b; Gluck and Thompson 1987; Hawkins 1989a, b; Hawkins and Kandel 1984; Morris 1990; Raymond et al. 1992). For example, Gingrich and Byrne (1987; Byrne and Gingrich 1989; Byrne et al. 1989) illustrated how a single-cell model that incorporated the activity-dependent neuromodulation learning rule can simulate several simple features of classical (Pavlovian) conditioning. Moreover, when incorporated into a simple three-cell network, the activity-dependent neuromodulation learning rule can simulate several higher-order features of classical conditioning, such blocking and second-order conditioning (Baxter et al. 1991; Buonomano et al. 1990; Byrne et al. 1991b; Gluck and Thompson 1987; Hawkins 1989a, b). Finally, when incorporated into a seven-cell network which spontaneously generates patterned activity, the activity-dependent neuromodulation learning rule can simulate many features of operant (instrumental) conditioning (Baxter et al. 1991; Byrne et al. 1991b; Raymond et al. 1992). Thus, it appears that a full understanding of the relationship between synaptic plasticity and learning requires an appreciation of different forms of learning and of the different neural architectures within which a given type of synaptic plasticity is embedded.

ACKNOWLEDGMENTS

We thank Dr. M. Mauk for his comments on an earlier version of this manuscript. This work was supported by Air Force Office of Scientific Research Grant 91-0027 and National Institute of Mental Health Award K02-MH00649.

Realistic Network Models of Distributed Processing in the Leech

Shawn R. Lockery and Terrence J. Sejnowski

The distributed nature of information processing in the brain presents a great challenge to systems neuroscience. Whether one considers the processing of sensory information or the control of motor responses, the output of the nervous system is a function of large populations of simultaneously active neurons. In all but the simplest reflexes, neurons are arranged in bewildering networks of parallel and feedback pathways, making a purely intuitive understanding of the system difficult if not impossible.

Neural models and computer simulations are approaches to this complexity that could help us achieve a deeper understanding of distributed processing. Useful predictions of a model must be directly related to the experiments that are performed. Since much of the current physiological effort in systems neuroscience is devoted to making single-unit recordings, models are needed that predict the responses of individual neurons. However, models sufficiently detailed to predict responses of individual neurons require a very large number of parameters, including the type and distribution of voltage-sensitive channels, the anatomical position of synapses, and the sign and strength of synaptic connections. Since the value of such parameters is generally not known, one approach is to study the effect on the model of systematically varying each parameter over its physiological range. In general, however, this is not an efficient strategy because there are often too many parameters and their physiological range is often quite large.

There are now a variety of network optimization algorithms that adjust parameters such as connection strengths in artificial neural networks (Hinton 1989). The algorithms adjust each parameter in the network to reduce the overall error in the performance of the network. These algorithms differ according to the type of error information available and the way it is used to change the parameters. Optimization techniques are being applied to network models of known biological circuits where the number of parameters is too large to be studied individually. The implications of such models for our understanding of biological networks has been difficult to establish because of the many simplifying assumptions involved in the first generation of neural network models

(Sejnowski et al. 1988). This chapter reviews our application of network training algorithms to a distributed processing system in a well-defined invertebrate preparation, the local bending reflex of the medicinal leech. The chapter starts with an overview of two common modeling strategies: compartmental and connectionist or neural network models. Its central point is that it is now possible to combine the compartmental and connectionist approaches to create realistic models with large numbers of parameters. Much of the chapter is devoted to exploring in detail a specific example of this combined approach using the local bending reflex as a model system. We first review the behavioral physiology of the local bending circuit and the conceptual issues that prompted our modeling efforts. The model is then presented in detail and its performance is evaluated in terms of a distributed processing hypothesis of the local bending reflex. We conclude by suggesting how additional physiological constraints could be added to predict the identity of as yet unidentified interneurons in the leech.

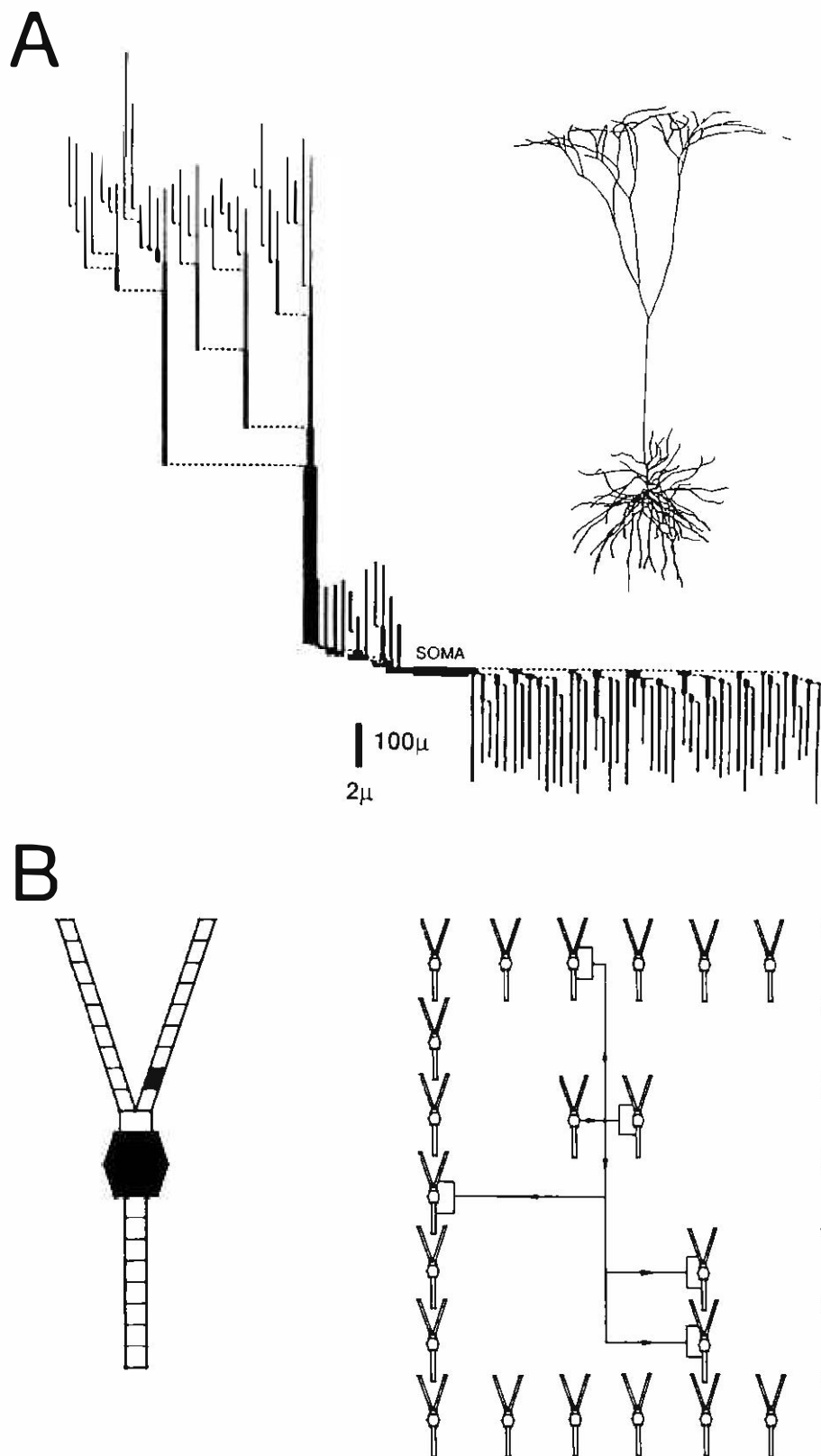
MODELING OVERVIEW

Compartmental Models

In some of the most detailed simulations produced to date, compartmental models are used to represent individual neurons. The morphology of each neuron is divided into a number of isopotential electrical compartments, with neighboring compartments representing nearby regions of the neuron, and currents pass between compartments via axial resistances (figure 5.1A). The voltage V in a passive compartment is given by:

$$\tau \frac{dV}{dt} = -V + RI_m \quad (5.1a)$$

where τ is the time constant of the compartment, R is the input resistance, and I_m is the net transmembrane current (for more details, see Segev et al. 1989). Action potentials are achieved by locating voltage-dependent ion channels in axonal and somatic compartments. Synapses are modeled by putting synaptic conductances in dendritic compartments. The collective behavior of a large network of neurons is then investigated by connecting the model neurons together according to the known circuitry of the brain region of interest (figure 5.1B). This strategy has been used to study the emergent properties of models that incorporate the cell types and connectivity of such systems as the hippocampus (Traub and Wong 1982; Traub et al. 1989), the olfactory bulb and piriform cortex (Bower 1990), and the neocortex (Bush and Douglas 1991). For example, Bush and Douglas (1991) showed that pyramidal cells with different intrinsic bursting frequencies in a model of visual cortex quickly synchronized their bursting as a result of global feedback



Compartmental Models

Figure 5.1 (A) Equivalent compartmental model of a pyramidal cell in visual cortex. Each rectangle is a single compartment whose length and width matches the length and diameter of the corresponding segment of the neuron (inset). Dotted lines indicate branch points in the dendritic tree. Length and width of compartments are shown on different scales, as the scale bar indicates. (Part A reprinted from Lytton and Sejnowski 1991, with permission of the American Physiological Society.) (B) Model of area CA3 of the hippocampus using a network of compartmental neurons. Each neuron in the network is a simplified compartmental neuron comprising a soma and three dendrites represented by a total of 28 compartments. The soma and one dendritic compartment (filled polygons) contain active conductances producing action potentials. Excitatory connections (arrows) in the network were assigned randomly according to the probability of connections among neighboring CA3 neurons. (Part B reprinted from Traub and Wong 1982, © AAAS.)

inhibition and intracortical reexcitation. Thus, synchrony between cortical neurons responding to coherent visual stimuli could be an emergent property of intracortical connectivity.

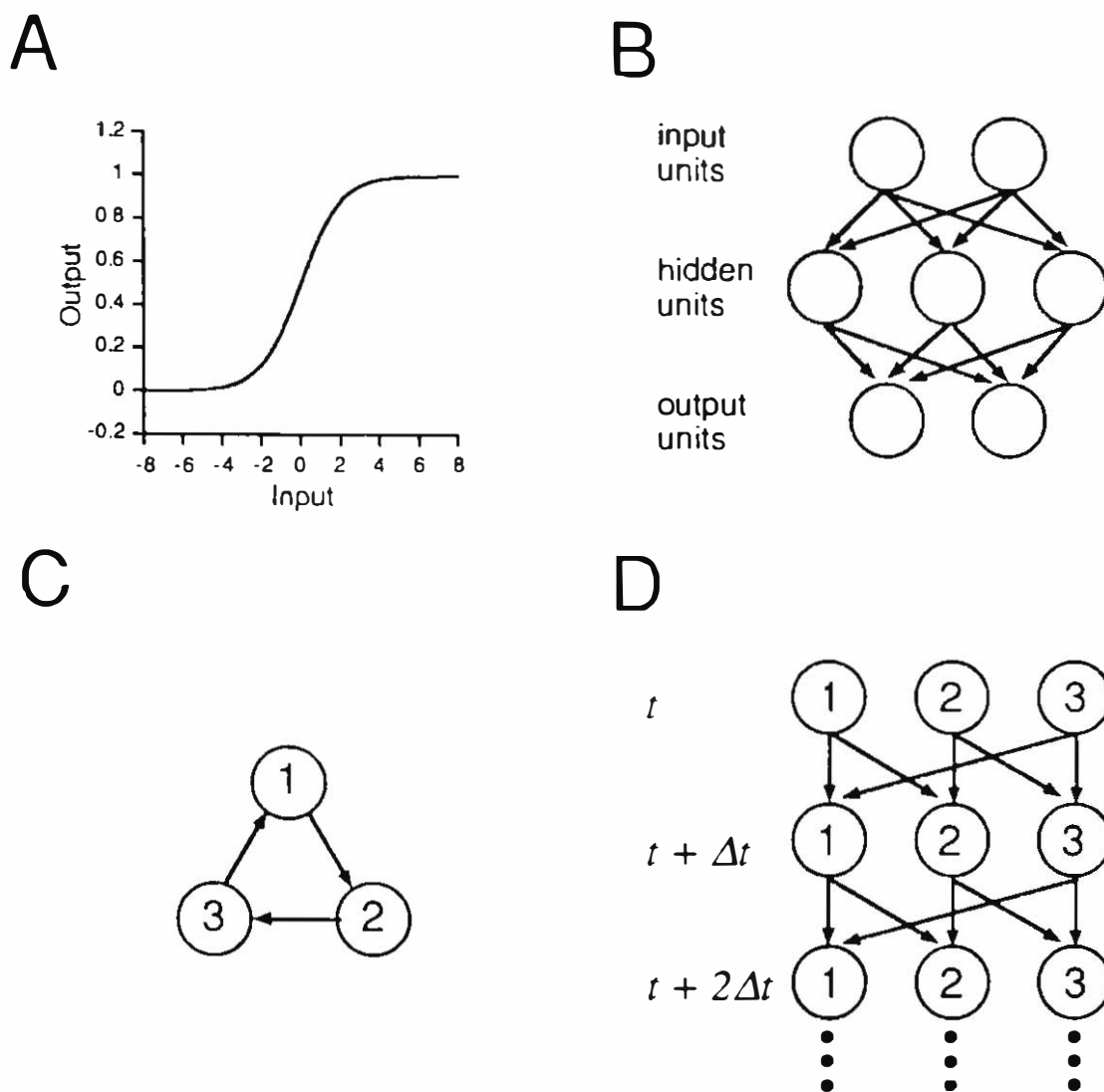
Neural Network Modeling

A neural network model trades the realism of compartmental models for control over the network's input-output function. Network models are composed of simplified neuronlike elements or *units* connected by synapses of various strengths or *weights* (see figure 3.2). Each unit has a state or activity level that is a function of its net input. The net input to a unit is the weighted sum of the activities of the units from which it receives a connection. The output of a unit is generally a nonlinear function of its net input or activity level. Many networks use a sigmoidal function (figure 5.2A). This approximates the relationship in real neurons between net input and firing frequency, since weak inputs produce little or no activity (*cutoff*), and strong inputs drive the unit to its maximum firing rate (*saturation*). Variations in the construction of network models include the use of binary units that are either on or off (Hopfield 1982), and units that multiply their inputs in addition to adding them (Durbin and Rumelhart 1989).

Model networks usually contain three types of units (figure 5.2B; see also figure 3.3): (1) *input units* that receive inputs only from sources external to the network; (2) *output units* that define the output state of the network; and (3) *hidden units* that receive inputs from sensory neurons and provide the inputs to the output units. A particular set of activation values on the input units constitutes an input vector and the set of activations on the output units constitutes the output vector. The input-output function of a network is the mapping or list of associations between input and output vectors, and it is a function of the weights in the network. The power of network models derives from the use of optimization algorithms that can adjust the weights to achieve almost any input-output function, a task that is next to impossible if attempted by hand in large networks.

Backpropagation

Algorithms for optimizing weights fall into three classes: *supervised* procedures in which the algorithm is given a list or *training set* of desired input and output vector pairs (Rumelhart et al. 1986a), *reinforcement* procedures in which each input vector is associated with a scalar output assessment instead of an output vector (Barto and Jordan 1987), and *unsupervised* procedures in which there is no assessment of network output. However, the distinction between supervised and unsupervised procedures is not absolute. Some ostensibly unsupervised algorithms turn out on closer inspection to contain an internal assessment of performance or monitoring. It is thus more useful to distinguish between



Aspects of Network Models

Figure 5.2 (A) The sigmoidal input-output function for the neuronlike units used in many network models. (B) Feedforward network. (C) Recurrent network. (D) The network in C represented as a multilayer, feedforward network in which layers correspond to successive points in time. Units connect between but not within layers. The recurrent network in the feedforward configuration can be trained by backpropagation in much the same way as a feedforward network—by sending the error backward from later to earlier points in time.

monitored and *unmonitored* systems as well as supervised and unsupervised (Churchland and Sejnowski 1992).

Backpropagation is a supervised, unmonitored optimization procedure for networks with one or more layers of hidden units. At the start of optimization, the weights are assigned small values at random. An input vector is presented, and activity flows forward to the output units, determining an output vector. An error measure is then calculated by comparing the output vector to the output specified in the training set. This is repeated for each input-output pair in the training set. The weights in the network are then changed or updated by small amounts based on the total error produced by this pass through the training set. Each pass through the training set and its associated weight update is called a *training cycle*. After many cycles, a set of weights is found that reproduces the desired input-output function.

In mathematical terms, backpropagation is a form of gradient descent.

Weights are updated in the direction of the error gradient, and backpropagation is a way to estimate the derivative of the total error with respect to each weight. The total error produced on any pass through the training set is a function of the weights. The derivative of this function with respect to each weight tells whether increasing the weight will increase or decrease the error. Since our objective is to find weights that minimize the error, we decrease the weights with positive derivatives, and increase weights with negative derivatives. It turns out that the derivative of the error with respect to each weight can be calculated explicitly from the activity of the presynaptic unit, together with the net input and error of the postsynaptic unit (Rumelhart et al. 1986a). The procedure for multilayer networks is called backpropagation because the error of the postsynaptic unit is calculated recursively in terms of the error of the units to which it projects. In practice, this involves sending the error back from the output units to the hidden units. (For additional commentary on backpropagation, see chapter 3.)

Recurrent Backpropagation

Training networks with recurrent or feedback connections provides a special challenge because the effect of a weight on the error is different at different times (figure 5.2C). Several closely related algorithms exist for adjusting connections in feedback networks (Pearlmutter 1989; Pineda 1987, 1989; Rumelhart et al. 1986a; Williams and Zipser 1990). The most intuitively clear version is called *backpropagation through time* (Williams and Zipser 1990). The idea is to convert the feedback network into an equivalent feedforward network by breaking its activity into discrete time steps (figure 5.2D). The activity of each neuron at each time step is represented as a single neuron in an equivalent feedforward network. The resulting network can then be optimized as if it were a feedforward network. This procedure is referred to as *recurrent backpropagation* to distinguish it from the original form of backpropagation (Rumelhart et al. 1986a) in which consecutive layers do not represent time steps. The original form is sometimes called *static backpropagation*.

Models of Real Circuits Optimized by Static Backpropagation

The first application of network models for real-life neurophysiological modeling arose from the insight that optimization algorithms originally developed for the study of cognitive processes could be used to train networks to reproduce physiologically relevant input-output functions (Lehky and Sejnowski 1988; Zipser and Andersen 1988a). Analysis of the response properties of hidden units in such models often gave new insight into the possible function of neurons in the biological network.

Zipser and Andersen (1988a) used a three-layered feedforward network model as a means of investigating the possible spatial processing performed by neurons in area 7a of parietal cortex. The network was

optimized by static backpropagation to compute the location of a visual stimulus in head-centered coordinates from the input units that signaled either the location of the stimulus on the retina or the angle of gaze of the eyes with respect to the head. After optimizing, the hidden units in the model network had complicated, nonlinear spatial receptive fields that were qualitatively indistinguishable from the receptive fields of area 7a neurons. The correspondence between units in the model and area 7a neurons provided a missing conceptual link between area 7a and spatial localization. In addition, it provided a new hypothesis concerning how the spatial receptive fields of area 7a neurons could participate in the representation of head-centered spatial information.

In a related study, Lehky and Sejnowski (1988) optimized a network to compute the three-dimensional shape of an object from the two-dimensional pattern of light and dark on the retina. The model showed that receptive fields that seem designed for edge detection arise in networks whose sole purpose is to determine three-dimensional shape. Their results pointed out that receptive field information in isolation is an unreliable predictor of the function of neurons in distributed processing systems. In later investigations (Lehky et al. 1992), the activity of complex cells in monkey striate cortex was recorded in response to hundreds of visual images. For each cell, a network model was optimized to reproduce the recorded responses when presented with the same stimuli. After optimization, the model gave accurate responses to novel stimuli, indicating it captured more of the input-output transfer function of the complex cell than previous analytical methods.

Anastasio and Robinson (1990) optimized a network model of the vestibulo-ocular reflex (VOR) to produce eye velocity signals appropriate to head velocity and stimulus-pursuit commands reaching the vestibular nucleus. After optimization, hidden units in the model received both head velocity and pursuit inputs, but to varying degrees, as has been observed experimentally. Their model showed that mixed velocity and pursuit inputs could be an integral part of a distributed vestibulo-oculomotor system.

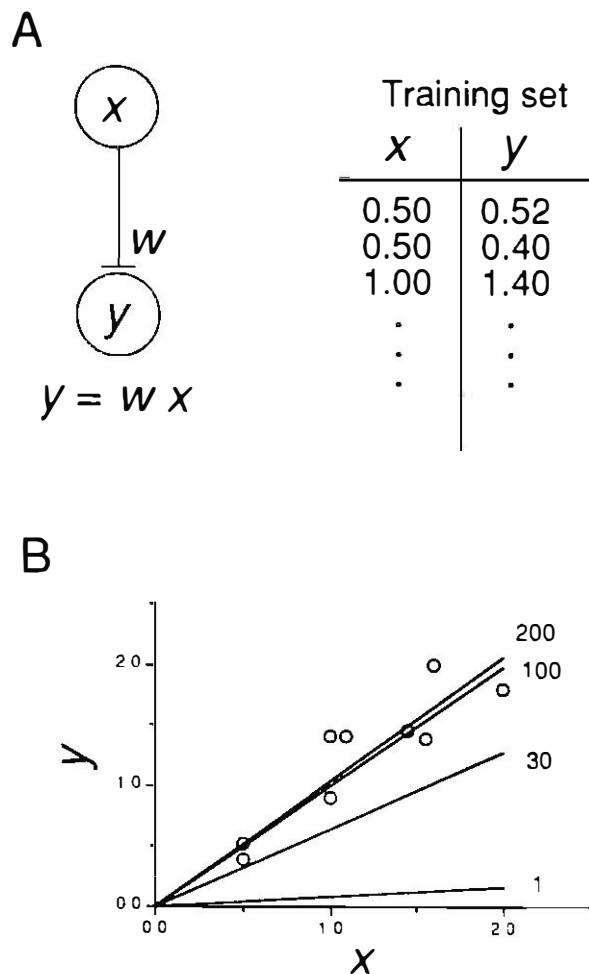
Exploiting the advantages of an invertebrate system, Lockery et al. (1989) were able to measure directly the input-output function of the local bending reflex in the leech, a simple withdrawal response, by stimulating sensory neurons and recording the resulting synaptic potentials in the motor neurons controlling the reflex. Their model incorporated biological constraints such as lateral inhibition between motor neurons, left-right symmetry among pairs of homologous interneurons, and solely excitatory sensory-to-interneuron connections. After optimization to the known input-output function, the receptive and projective fields of model interneurons closely matched those of identified local bending interneurons. The model showed that local bends in response to dorsal, ventral, or lateral stimuli could be produced by a single population of multifunctional interneurons. More recently Wittenberg

and Kristan (1992) have extended the model to include a second withdrawal behavior—shortening. Their results suggest that at least two new types of interneuron involved in leech withdrawal behaviors remain to be found.

Models of Real Circuits Optimized by Recurrent Backpropagation

Static backpropagation is sufficient for problems in which the biological network is, for all intents and purposes, a feedforward network. However, the great majority of neural networks possess feedback connections. Feedback connections require the use of recurrent backpropagation because such connections act over time. One place where feedback connections play a prominent role is in central pattern generators. Pearlmutter (1989) and Doya and Yoshizawa (1989) showed how recurrent backpropagation can be applied to idealized central pattern generators in which each neuron is a passive summing node. Rowat and Selverston (1991) extended this approach for specific cases like the gastric mill central pattern generator of the lobster, where the neurons have active conductances and electrical synapses. Lockery et al. (1990) applied recurrent backpropagation to a model of the leech local bending reflex that included two important features omitted from the previous static model (Lockery et al. 1989): feedback connections and the time course of physiologically recorded synaptic potentials. The recurrent backpropagation model of local bending is described in detail below. The numerous feedback connections in the vestibulo-ocular reflex have prompted a number of models using recurrent backpropagation. Anastasio (1991a) showed that inhibitory feedback connections among vestibular neurons in the brain stem could be the site of integration of the head-velocity signal in the VOR. This model also reproduces the recovery of the response properties of vestibular neurons following ablation of afferent inputs (Anastasio 1992), and suggests sites of plasticity underlying the recovery. Recurrent backpropagation has also been used to construct plausible models of the complex firing patterns exhibited by cortical neurons during the performance of various memory (Zipser 1991) and motor tasks (see chapter 7). A further model considers the possible role of basal ganglia in switching between motor tasks (Brotchie et al. 1991).

Two other groups have constructed dynamic models of the VOR, but used optimization strategies other than recurrent backpropagation. Arnold and Robinson (1990) used numerical differentiation. In this technique, the derivative of the error with respect to a given weight is estimated numerically by increasing the weight slightly and measuring the amount by which error increases or decreases. After doing this for every weight, the weights are updated by decreasing weights with positive derivatives and increasing weights with negative derivatives. Quinn et al. (1992) used a global optimization procedure in which minima of the error function are found by finding the roots of polynomial fits to the error function (Bremermann 1970). Both of these procedures are computationally



The Analogy Between Backpropagation and Curve Fitting

Figure 5.3 (A) On the left is a network with one input unit, one output unit, and one weight. The activity of the output unit y is the product of the weight w and the activity of the input unit x . In backpropagation, the task is to find the value of w that best reproduces the training data shown on the right. Note that x and y are used differently in chapter 3. (B) Optimization of w by backpropagation. The circles are the data points in the training set. The number of iterations of the training set is shown to the right of each w shown. A random initial value of w gives a poor fit to the data (1 iteration). The fit improves with each cycle of optimization (30, 100, and 200 iterations shown).

more expensive than backpropagation and so work best for relatively small networks.

Backpropagation and Biological Plausibility

The historical tendency to speak in terms of training when describing the process of adjusting connection strengths via backpropagation or other optimization algorithms has lead some to question the relevance of network models to biology. In fact, none of the examples above used backpropagation as a model for the process by which connection strengths are established in the animal during development (ontogenesis) or during evolution (phylogenesis). The interest is strictly in the end point of the optimization process. In mathematical terms, the task of optimizing a model network to reproduce a desired input-output function is identical to a curve-fitting procedure in which the weights are the coefficients or parameters to be adjusted, and the input-output relations in the training set are the data to be fit. This point is illustrated in figure 5.3. Suppose

we wished to fit a line starting at the origin through the data points in figure 5.3B. In neural network terms, this amounts to supposing that the x coordinate of each point is the activation of an input unit and the y coordinate is the desired activation of an output unit (figure 5.3A). If the input-output function of the output unit is linear, then $y = wx$, and we wish to find the w that minimizes the error in predicting for each x the corresponding y in the data set. A random initial value of w gives a poor fit, but after 200 cycles through data, a satisfactory fit is achieved. Backpropagation is best viewed as one of many techniques for finding an optimum set of parameters when fitting a function. The issue of biological plausibility is moot as far as optimization is concerned. For example, linear regression is not biologically plausible, yet it is a perfectly legitimate means of fitting functions to biological data. This having been said, approximations to backpropagation that have greater biological plausibility have been studied (Mazzoni et al. 1991), and some undesigned-for similarities between backpropagation and neuronal networks are discussed in chapter 3.

Combinations of Compartmental and Neural Network Models

With the introduction of recurrent backpropagation it is possible to combine the strengths of compartmental and neural network models. In a recurrent network model, the state of each unit is its activity level. However, if the state is taken instead to represent a voltage, then each unit in the network model is formally equivalent to an electrical compartment in a compartmental model. In such a model, units communicate by passing currents rather than weighted activations, and the sigmoidal input-output function represents the relationship between presynaptic voltage and postsynaptic current. It thus becomes possible to use neural network training algorithms such as backpropagation to adjust connections between compartments in realistic models (Rowat and Selverston 1991).

A CASE STUDY: THE LEECH LOCAL BENDING REFLEX

Background

The local bending reflex of the leech proved to be an ideal preparation in which to apply the combined compartmental and network modeling approach to a real biological system. First, the reflex is subserved by a comparatively small number of sensory neurons, interneurons, and motor neurons, each of which is re-identifiable from animal to animal. Second, the input-output function of the reflex could be measured with unusual precision by pairwise intracellular recordings from sensory neuron and motor neurons, providing a more realistic training set than is normally available. Third, interneurons in the reflex receive multiple sensory inputs and form a distributed representation of stimulus location, mak-

ing it difficult to construct a model by hand. Finally, physiological estimates of the connection strengths of interneuron inputs and outputs could be made. This allowed explicit comparisons between the connectivity of model and actual interneurons.

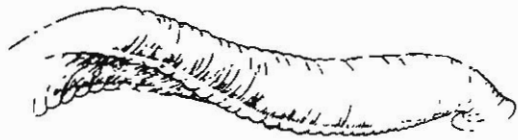
Local bending is a directed escape reflex (figure 5.4A). In response to a moderate mechanical stimulus, the leech withdraws from the site of contact. This is accomplished by contracting longitudinal muscles beneath the stimulus, and relaxing longitudinal muscles on the opposite side of the body, resulting in a U-shaped local bend. Thus, a dorsal stimulus causes a dorsal bend, a ventral stimulus a ventral bend, and a lateral stimulus a lateral bend. The local bending reflex can be elicited in a semi-intact preparation consisting of a single midbody ganglion and the associated segment of body wall. It has thus been possible to identify the sensory neurons, motor neurons, and a subpopulation of the interneurons contributing to the reflex (Kristan 1982; Lockery and Kristan 1990b; Nicholls and Baylor 1968). Each ganglion contains sensory neurons responding to touch and pressure (Nicholls and Baylor 1968). Three pairs of T cells respond selectively to touch on the dorsal (TD), ventral (TV), and lateral (TL) body surface. In addition, there are two pairs of P cells, which respond to pressure on the dorsal (PD) or ventral (PV) body surface, respectively. Several types of motor neurons innervate the longitudinal muscles that control the reflex (Ort et al. 1974; Sawada et al. 1976; Stuart 1970). These include excitors and inhibitors of the dorsal longitudinal muscles (DEs and DIs), excitors and inhibitors of the ventral longitudinal muscles (VEs and VIs), an excitor of lateral longitudinal muscle, and the L cell, which has a combined dorsal and ventral field. The innervation of each motor neuron type is restricted to a single side of the body (left or right).

As a first step in understanding how local bending is produced, we focused on a representative subset of the sensory and motor neurons contributing to local bending (figure 5.4B). We used activation of P cells as sensory input to the reflex since dorsal, ventral, and lateral stimuli each activate at least one P cell, and the effect of P cells is greater than that of the T cells (Kristan 1982). Motor outputs were monitored by recording from DEs, DIs, VEs, and VIs. The lateral motor neuron was not monitored in these initial investigations because, for technical reasons, it is somewhat more difficult to locate and identify. The L cell was not monitored in this initial study because its response is much briefer than the response of the other motor neurons (Kristan 1982).

The simplified local bending network retains essential features of the complete network. The spatial location of sensory input is encoded on an array of sensory neurons, and motor output on an array of motor neurons, which are highly interconnected by both chemical and electrical synapses. The connection between sensory neurons and motor neurons is mediated by a layer of interneurons. In computational terms, the task of the interneurons is to compute the input-output function of the

A

Resting



Dorsal



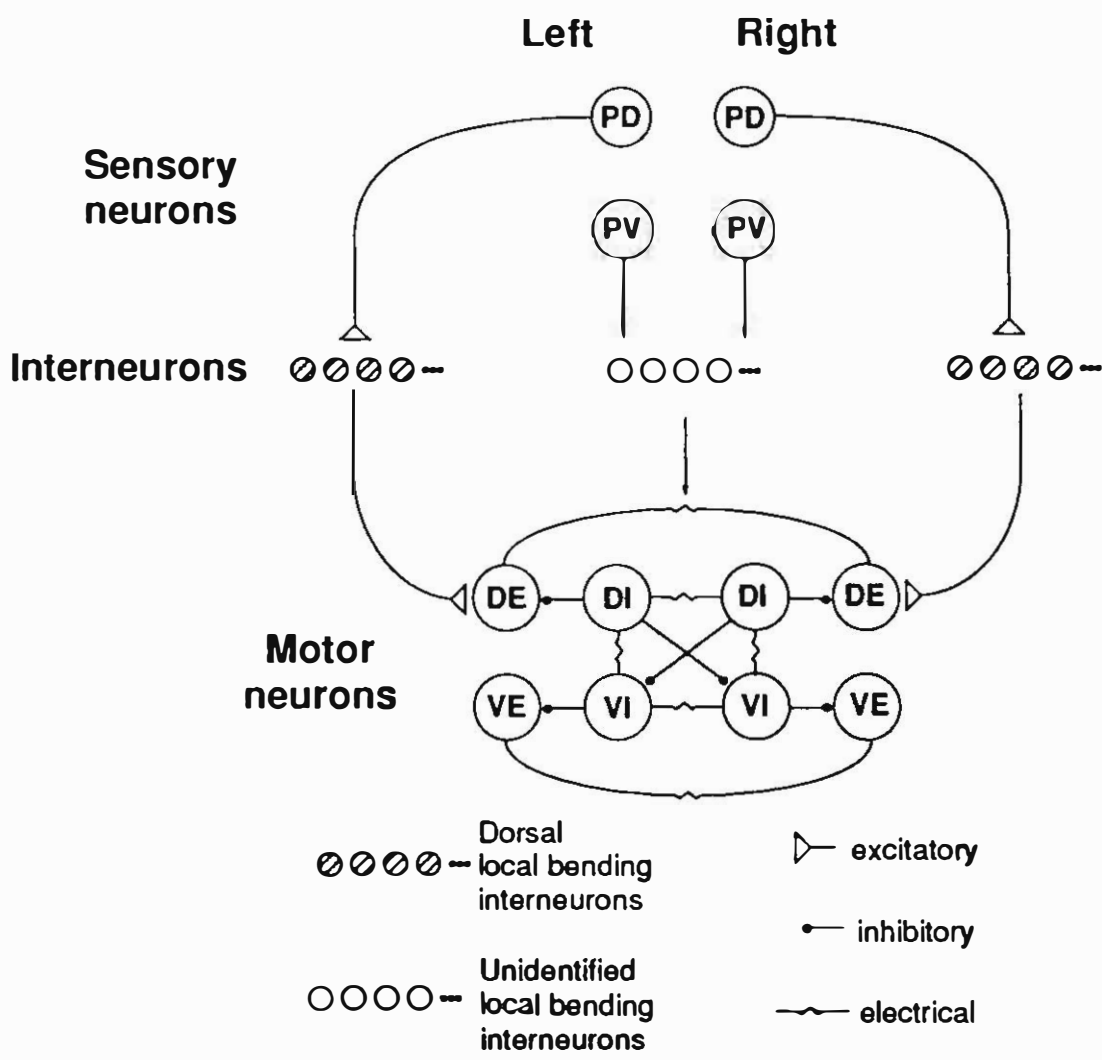
Ventral



Lateral



B



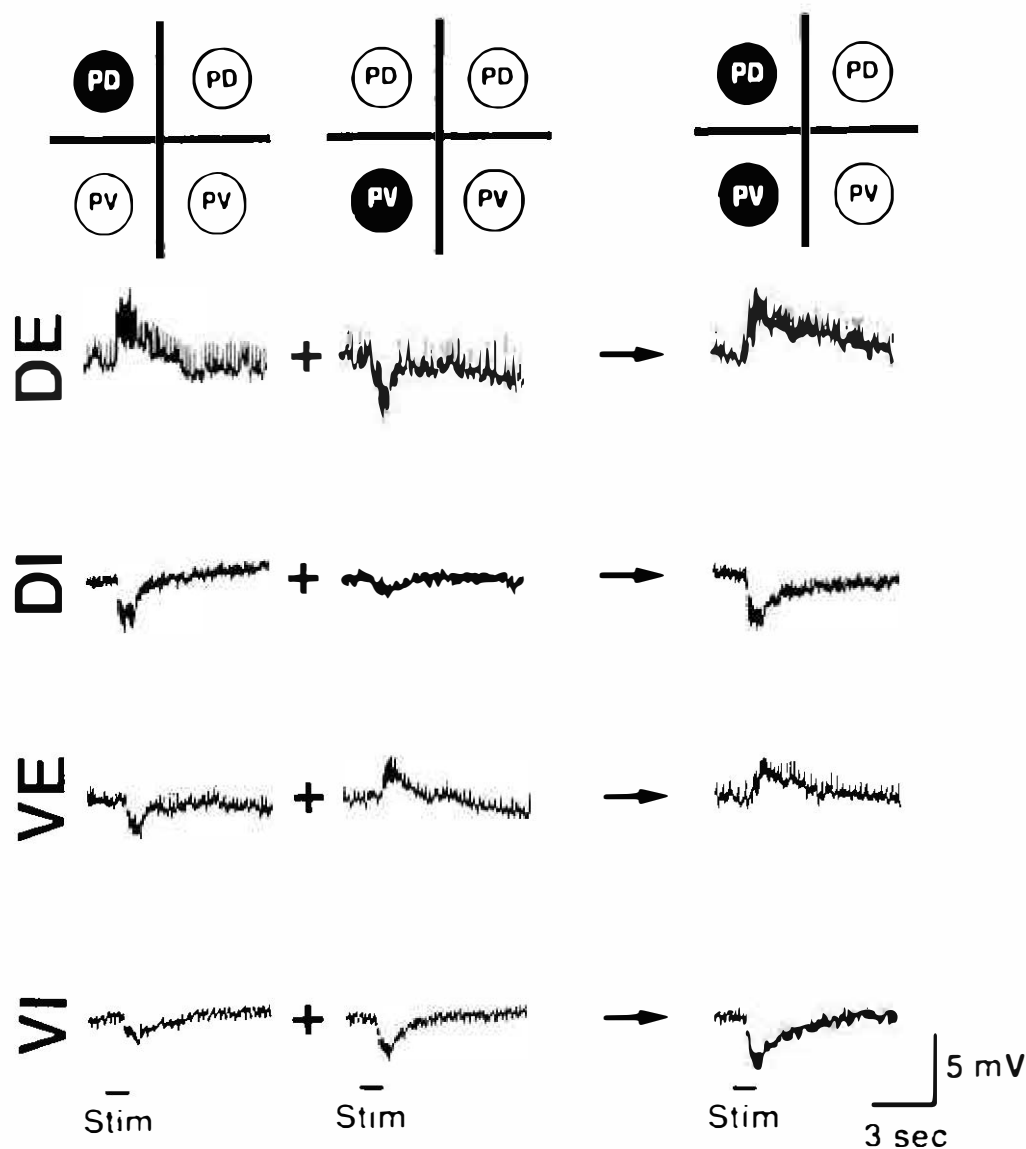
reflex: the mapping between each pattern of P cell stimulation and the pattern of motor neuron excitation and inhibition sufficient to withdraw from the stimulus. Synaptic transmission from sensory neurons to interneurons involves action potentials in the presynaptic neuron (Lockery and Kristan 1990b), but transmission from interneurons to motor neurons (and between motor neurons) does not; as in the vertebrate retina (Roberts and Bush 1981), synaptic activation is a graded function of presynaptic voltage.

Measuring the Input-Output Function

We began our analysis with the input-output function of the local bending reflex (figure 5.5; Lockery and Kristan 1990a). Using a standard stimulus sufficient to elicit local bending responses, P cells were activated, in pairs or individually, while intracellular recordings were made from all eight motor neurons in the simplified network. We examined in detail the motor neuron synaptic potentials produced by eight patterns of sensory input. These included four pairs of P cell stimulation (the left and right PD, the left and right PV, left PD and PV, right PD and PV), as well as each P cell stimulated individually. In general, the response of the network to stimulation of pairs of P cells was well predicted by the sum of the responses to the same two P cells stimulated individually. For example, the motor neuron VI, in response to ipsilateral stimulation of PV and PD, receives an IPSP that is approximately the linear sum of the IPSP produced by individual stimulation of PV and PD. However, there were exceptions to this, as in the case of the ipsilateral VE, which receives an EPSP in response to ipsilateral stimulation of PV and PD, despite the fact that the individual effects of PV and PD tend to cancel each other out. In this case, the nonlinearity of the response to the paired P cell pattern is in the direction that is behaviorally efficacious, since increased excitation of VE promotes withdrawal from the site of stimulation. Similar results were found for other patterns of paired P cell stimulation. Thus, the basic patterns of motor output underlying local bending behavior are produced by the linear combination of individual

Local Bending in the Leech: Behavior and Circuitry

Figure 5.4 (A) Local bending behavior. The leech withdraws from the site of contact by contracting longitudinal muscles beneath the stimulus, and relaxing longitudinal muscles on the opposite side of the body. (B) Simplified local bending circuit. The strongest input to the reflex is provided by pressure sensitive mechanoreceptors, the P cells, which have dorsal (PD) or ventral (PV) receptive fields. Behavior is produced, in part, by dorsal and ventral excitatory (DE, VE) and inhibitory (DI, VI) motor neurons. The contributions of longitudinal muscle motor neurons with lateral projective fields, as well as the motor neurons controlling circular and oblique muscles, have not been studied in detail. Input from sensory to motor neurons is mediated by a layer of interneurons, a subpopulation of which have been identified.



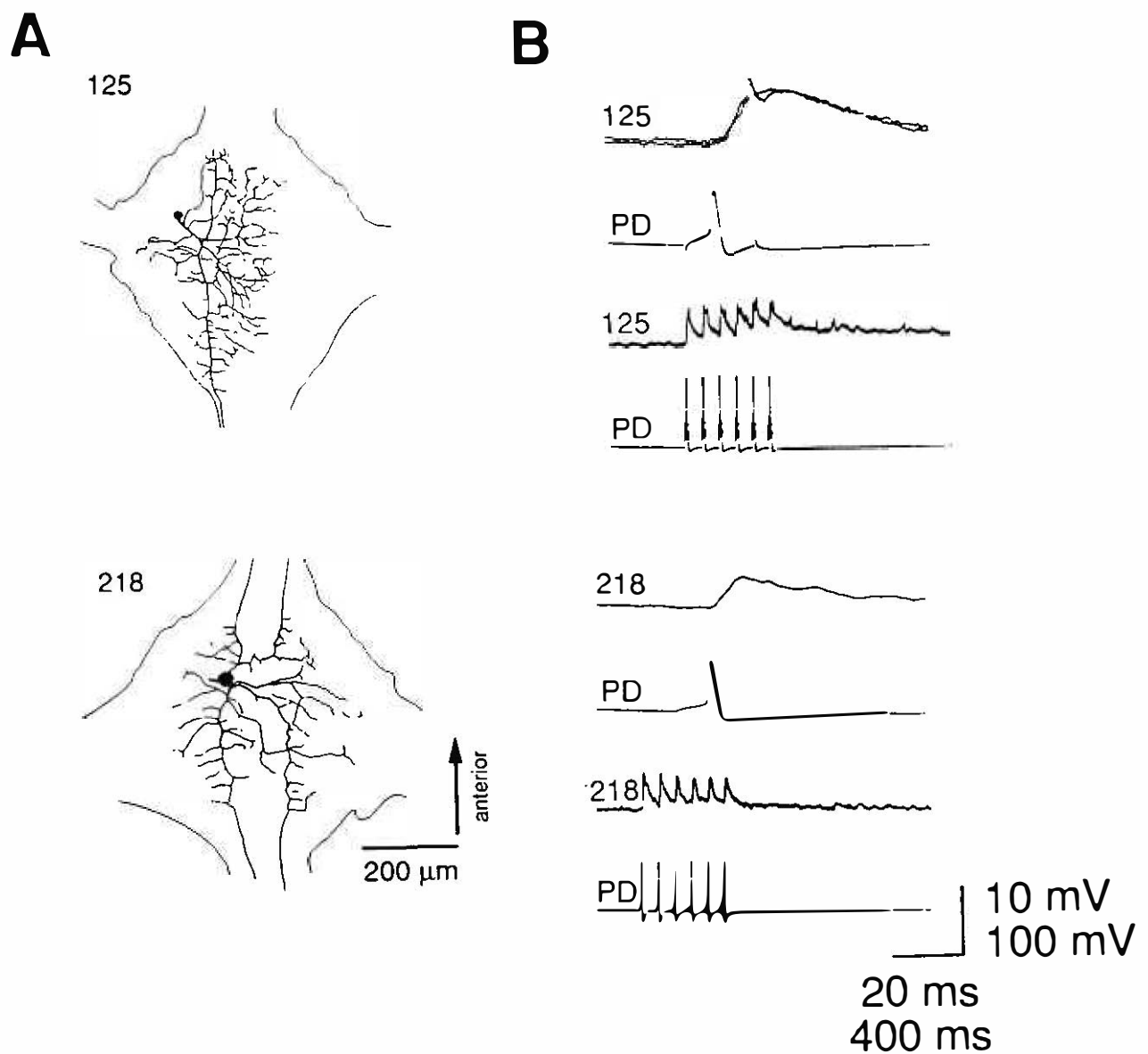
Local Bending Input-Output Function

Figure 5.5 Synaptic potentials recorded intracellularly from four motor neurons in response to single and paired P cell stimulation. The filled circles at the top of each column indicate which P cells were stimulated. Each row shows the synaptic potential recorded in the indicated motor neuron. Motor neurons have projective fields ipsilateral to the sensory neurons; recordings from contralateral motor neurons were also made but are not shown. The duration of the stimulus is shown by the bars below. All recordings are from the same preparation. Similar experiments were made for all possible patterns of paired P cell stimulation. (Reprinted from Lockery and Kristan 1990a by permission of the *Journal of Neuroscience*.)

P cell effects, with minor enhancements due to a nonlinearity in the system (Lockery 1989).

Identification of Interneurons

As a working hypothesis of sensorimotor integration in the reflex, we adopted a *dedicated interneuron* model in which each form of local bending is under the control of a single interneuron, or type of interneuron. We imagined that dorsal local bending is produced by one or more specific, dorsal bending interneurons that are excited by dorsal P cells and receive no input from (or are inhibited by) ventral P cells. The unique pattern of excitation and inhibition of longitudinal muscle motor neurons is determined by the pattern of excitatory and inhibitory con-



Identifying Properties of Local Bending Interneurons

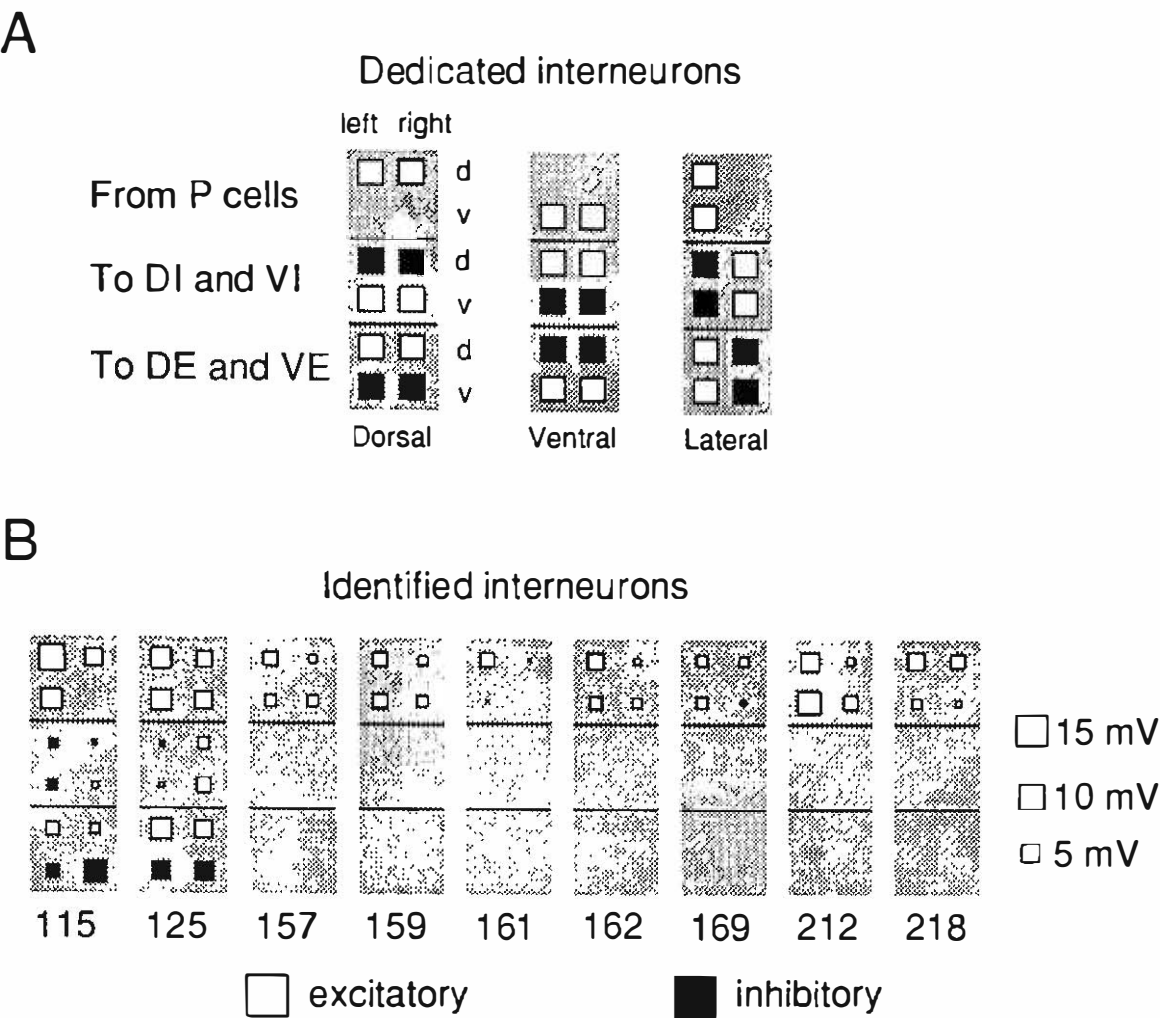
Figure 5.6 (A) Morphology. Camera lucida drawings were made from whole mounts of cells injected ionophoretically with Lucifer yellow. Identifying numbers were assigned according to the most common location of their cell body on a standard map of the ganglion (Muller et al. 1981). (B) Physiological properties. Within each panel, the top trace shows the response of the interneuron to the single PD action potential shown in the second trace. The third trace shows the response to the train of PD impulses shown in the fourth trace. In some traces, two or three sweeps (several seconds apart) are superimposed to show the extent to which latencies were consistent. Vertical calibration: first and third traces, 10 mV; second and fourth traces, 100 mV. Horizontal calibration: first and third traces, 20 ms; second and fourth traces, 400 ms. (Reprinted from Lockery and Kristan 1990b by permission of the *Journal of Neuroscience*.)

nections from dorsal bending interneurons to motor neurons. Similar types of interneurons are specific for ventral and lateral bends. Because one or the other of the P cells stimulated in lateral local bending is also stimulated in dorsal or ventral bending, the lateral bending interneurons must be capable of overcoming the effects of the dorsal and ventral bending interneurons. This could be accomplished by inhibitory connections from the lateral local bending interneurons onto other interneurons.

To determine whether local bending is produced by dedicated interneurons of the type we hypothesized, we sought to identify interneu-

rons in the leech contributing to local bending and to determine their sensory and motor neuron connections. As a first step, we searched for interneurons contributing to dorsal local bending; that is, interneurons that received excitatory input from at least one dorsal P cell, and in turn excited at least one DE motor neuron. Interneurons were impaled and a dorsal P cell was stimulated to fire a train of action potentials sufficient to elicit local bending (10 Hz, 0.5 s). Any neuron that received an excitatory synaptic potential from the dorsal P cell was then depolarized by current injection while action potentials of the DE motor neuron were recorded extracellularly. Neurons that were excited by the dorsal P cell and excited DE were considered dorsal bending interneurons. Each was filled with a fluorescent dye and its morphology was recorded. A survey of more than 70 ganglia yielded a sample of 235 dorsal bending interneurons. On the basis of morphological and physiological criteria (figure 5.6), nine types of dorsal bending interneuron were repeatably identifiable from preparation to preparation. Eight of these were bilaterally paired neurons and one was an unpaired cell; thus, each ganglion contained 17 dorsal bending interneurons. Each interneuron sent at least one axon into the interganglionic connective (figure 5.6A), very likely projecting to the adjacent ganglion where it may make both input and output connections (Wittenberg and Kristan 1992). As a group, the 17 interneurons send 16 axons into the anterior connective and 12 axons to the posterior connective. Thus, each ganglion potentially contains functional processes of at least 28 additional local bending interneurons, bringing the number of local bending interneurons to 45. If the axons of local bending interneurons extend more than one ganglion, like other interneurons in the leech (Brodfuehrer and Friesen 1986), the number rises to 73 or more. Since it seems likely that a search for interneurons contributing principally to ventral or lateral bends would identify other interneurons, this number is expected to go even higher.

Each morphologically defined type of local bending interneuron was associated with a unique intracellular recording in response to P cell stimulation (figure 5.6B). Thus, interneurons differed systematically in the size and shape of the synaptic potential produced by single P cell action potentials, the presence or absence of P cell evoked action potentials, as well as in the presence of a fast or slow component in the response to trains of P cell action potentials sufficient to evoke the behavior. The presence or absence of action potentials is interesting because it suggests that the effects of some interneurons are mostly local, given the P cell stimulus intensities used, while others might have both local and distal effects. Similarly, the difference in fast and slow responses to trains of P cell action potentials suggests that longer sensory stimuli might select a different subset of interneurons than are selected by short stimuli, since the response of interneurons with substantial slow components will be selectively enhanced. In principle, therefore, the heterogeneity among interneurons in their response to



Connectivity of Hypothesized and Identified Local Bending Interneurons

Figure 5.7 (A) Hypothesized dedicated interneurons. Gray dominoes show the input and output connections of a single interneuron. The upper third of each shows connections from sensory neurons to the interneuron; the middle third shows the connections from the interneuron to the inhibitory motor neurons; the bottom third shows the connection from the interneuron to the excitatory motor neurons. White boxes indicate excitatory connections; black boxes indicate inhibitory connections. Interneurons specific for dorsal, ventral, and lateral are shown. (B) Identified dorsal local bending interneurons. Symbols are as in A, except that box area is proportional to the amplitude of the connection determined from intracellular recordings of interneurons or motor neurons. Blank spaces denote connections whose strength has not been determined.

sensory input could be the basis for differences in the spatial and temporal patterns of motor output as a function of stimulus intensity and duration, thus increasing the computational repertoire of the network.

The dedicated interneuron model predicts that dorsal bending interneurons will have input and output connections like those in figure 5.7A. To determine if this was the case, we measured the strength of the connections from each P cell to a given interneuron by activating each P cell individually with a standard intracellular stimulus and noting the peak amplitude of the synaptic potential in the interneuron. The strength of output connections to motor neurons was measured in a similar fashion. However, we were hampered by the fact that the motor neurons lie on the dorsal surface of the ganglion, while all but two of the interneurons lie on the ventral surface. At present, therefore, output connections are known only for cells 115 and 125. The connectivity of

the dorsal bending interneurons was consistent with the dedicated interneuron model in that most interneurons received substantial excitatory input from both dorsal P cells, and cells 115 and 125 excited both DEs and inhibited both VEs (figure 5.7B). However, the results were inconsistent with the dedicated interneuron model in two respects. First, with one exception, each interneuron received substantial input from at least one ventral P cell. This means that these interneurons are active during ventral local bends, when their excitatory effects on DEs are inappropriate. This also means these interneurons are active during lateral local bending, when the effect on the contralateral DE is inappropriate. Second, the effect of cells 115 and 125 on the inhibitory interneurons was not always the opposite of the effect on the excitatory motor neuron with the same projective field; for example, cell 125 excites both DE and DI on the right side.

Motivation for the Model

That most local bending interneurons are active in dorsal, ventral, and lateral local bends suggested that the simple, dedicated interneuron model is incorrect. Rather, the interneurons appear to form a distributed representation of sensory input. In the distributed model, each form of the local bending behavior is produced by the simultaneous action of many interneurons, whose combined outputs produce the associated motor output pattern. The distributed model thus has the advantage that it can account for ventral inputs to dorsal bending interneurons (figure 5.7B) and the conflicts between interneuron outputs to motor neurons. In this view of the network, connections of an interneuron that are inconsistent with one motor output pattern are offset by behaviorally consistent connections from other interneurons.

Although the distributed processing hypothesis was attractive, it remained to be shown whether it was consistent with the physiological details of the biological network. Specifically, was the distributed model consistent with the many lateral connections among the motor neurons, with the amplitude and time course of motor neuron synaptic potentials in response of all different patterns of sensory input, and with the type of interneurons identified?

Architecture of the Model

To test the distributed processing hypothesis, we constructed a model of the reflex and used backpropagation to adjust connection strengths between neurons (Lockery et al. 1990). The model comprised four sensory neurons (P cells), up to 40 interneurons, and eight motor neurons (figure 5.8A). Sensory neurons responding to touch (T cells; Nicholls and Baylor 1968) were excluded from this preliminary model, since their effects on behavior are small in comparison to the effects of the P cells (Kristan 1982). The number of interneurons represented, in round num-

bers, an upper estimate of total number of local bending interneurons, based on the number of local bending interneurons identified to date within a single ganglion, and the number of axons from local bending interneurons projecting into the ganglion from adjacent ganglia (Lockery and Kristan 1990b). Each of the eight motor neurons in the model represented one of eight types of longitudinal muscle motor neurons, in which there are two to four neurons per type (Ort et al. 1974; Stuart 1970). These comprise the excitors and inhibitors of dorsal longitudinal muscle, DE and DI, respectively, and the excitors and inhibitors of ventral longitudinal muscle, VE and VI, respectively.

All known chemical and electrical synaptic connections between motor neurons were included and connection strengths were determined from previous physiological recordings (Granzow et al. 1985; see below). Weights of feedforward connections from sensory neurons to interneurons, and from interneurons to motor neurons were adjusted by the backpropagation algorithm (Rumelhart et al. 1986a) so that the model reproduced the amplitude and time course of motor neuron synaptic potentials recorded in response to single and paired P cell stimulation in eight different patterns (Lockery and Kristan 1990a). The training procedure was not allowed to insert connections between interneurons, since functional connections of this type have not been found; nor was the training procedure allowed to insert feedback connections from motor neurons to interneurons, since only one such connection had been identified (Friesen 1989b).

The present model is more realistic than the previous model (Lockery et al. 1989), which excluded electrical and chemical synapses that would introduce feedback among the motor neurons. Feedback leads to networks whose activity evolves in time. Thus recurrent backpropagation (Pearlmutter 1989) was used instead of static backpropagation (Rumelhart et al. 1986a) to train output units (motor neurons) in the network to follow a predetermined time course. The present model is also more realistic because the response of model motor neurons matches the real motor neurons not only in amplitude, but also in time course. By incorporating temporal dynamics, the model can predict the time course of synaptic potentials in the interneurons. (An additional discussion of dynamic recurrent networks is provided in chapter 7.)

Model Neurons

Neurons were modeled as passive, single electrical compartments having in parallel an input resistance R and capacitance C to ground. In later models, multicompartmental neurons could be used, but one compartment was sufficient for the theoretical issues raised by the physiological data currently available. The rate of change of membrane potential (dV/dt) was determined by the present voltage and the sum of the

current introduced by synaptic inputs using equation 5.1a, indexed by each neuron i :

$$\tau_i \frac{dV_i}{dt} = -V_i + RI_i \quad (5.1b)$$

where τ_i is the neuronal time constant, and I_i is the sum of the current introduced by chemical and electrical synapses. The use of single rather than multicompartmental model neurons is a simplifying assumption based on the assumption that the interneurons are nearly isopotential during the 500 ms long stimulus trains used in physiological experiments (Lockery and Kristan 1990a).

Trains of action potentials in sensory neurons were modeled as stepwise increases in presynaptic voltage. To reduce computation, action potentials within the trains were not represented individually, since the duration of the action potential is short with respect to the time constants for chemical synaptic transmission. The amplitude of the step was scaled to keep the range of synaptic potentials in the interneurons within the range observed in intracellular recordings (Lockery and Kristan 1990b).

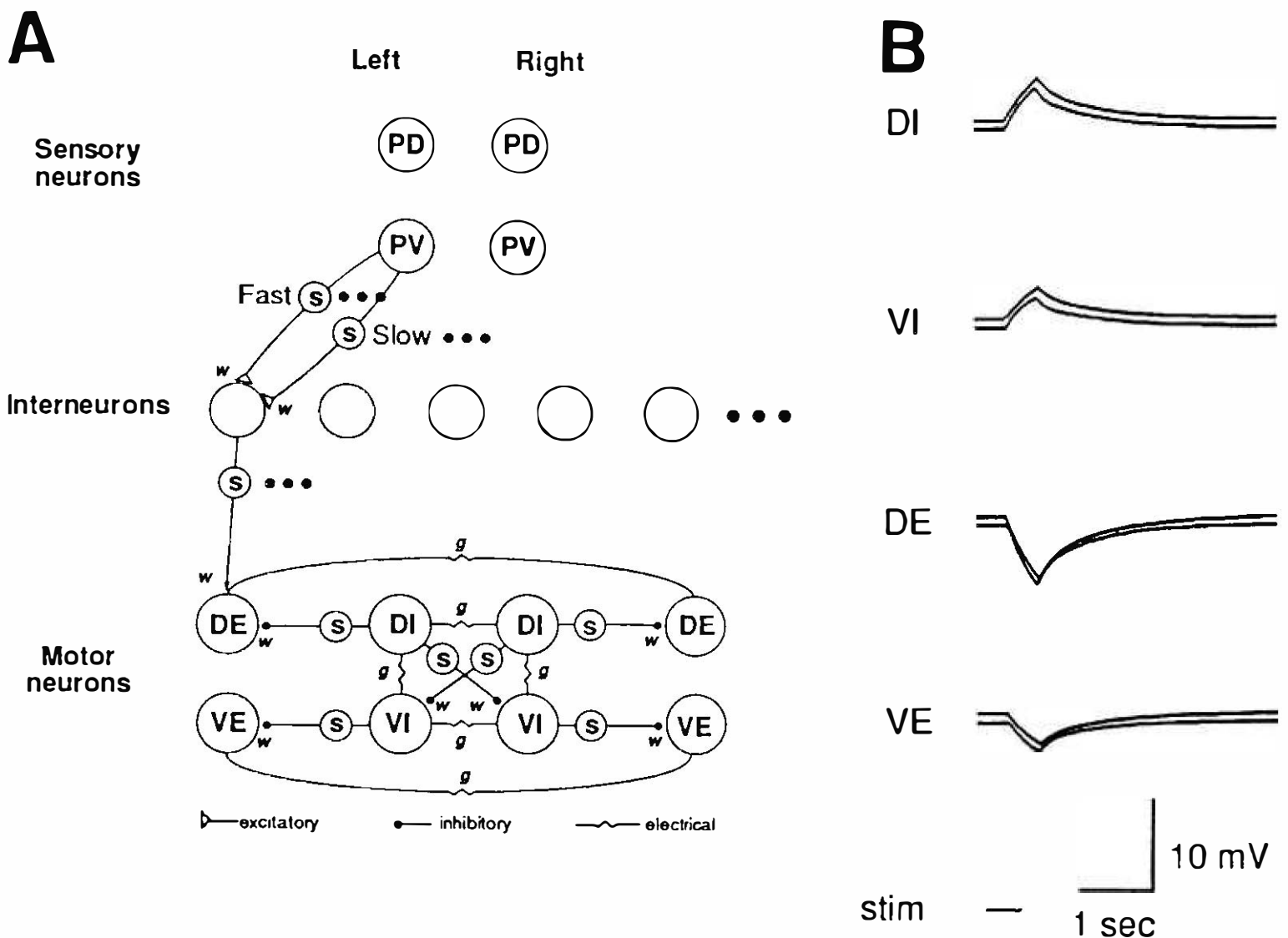
Synapses

Electrical synapses were modeled as ohmic conductances. Current introduced by electrical synapses was given by Ohm's law, and summed linearly according to the equation:

$$I_j^{elec} = \sum_i g_{ji}(V_i - V_j) \quad (5.2)$$

where g_{ji} is the conductance of the electrical synapse between neurons i and j . Within the ganglion, chemical synaptic transmission between inhibitory and excitatory motor neurons (Granzow et al. 1985), as well as from interneurons to other interneurons in the leech (Angstadt and Calabrese 1991; Friesen 1985; Thompson and Stent 1976), is a graded function of presynaptic voltage; it does not require the occurrence of action potentials, nor is it substantially affected by them. The synapse between inhibitory and excitatory motor neurons was used as a model for all chemical synapses in the network, since this connection has been most intensively studied. In physiological experiments (Granzow et al. 1985) a stepwise increase in presynaptic voltage in DI produced a slowly rising postsynaptic potential in DE. To account for the long synaptic rise time, synapse units (s-units, figure 5.8A) were inserted between pairs of neurons connected by chemical synapses. These can be viewed as additional compartments in which the state variable is synaptic activation rather than voltage. The activation of each s-unit S_{ji} was given by:

$$\tau_{ji} \frac{dS_{ji}}{dt} = -S_{ji} + f(V_i) \quad (5.3)$$



Network Model of the Local Bending Circuit

Figure 5.8 (A) Four sensory neurons were connected to eight motor neurons via a layer of 10 interneurons. Neurons were represented as single electrical compartments whose voltage varied as a function of time. Known electrical and chemical connections among motor neurons were assigned fixed connection strengths (g and w in the motor layer) determined from intracellular recordings. Interneuron input and output connections were adjusted by recurrent backpropagation. Chemical synaptic delays were implemented by inserting s-units between chemically connected pairs of neurons. S-units with different time constants were inserted between sensory and interneurons to account for fast and slow components of synaptic potentials recorded in interneurons. (B) Response of the model network to stimulation of the left ventral P cell after synaptic connections had been optimized using backpropagation. Each tracing shows the simulated intracellular recording from a single motor neuron, together with the target response from the data set against which connections in the model were optimized. To facilitate comparison of model and target synaptic potentials, the latter have been shifted upward by 1 mV. (Reprinted from Lockery and Sejnowski 1992 by permission of the *Journal of Neuroscience*.)

where τ_{ji} is the synaptic time constant and $f(V_i)$ is a sigmoidal function ($0 \leq f(V_i) \leq 1$) relating presynaptic and postsynaptic membrane potential.

The s-unit time constants combined the temporal dynamics of all the steps in transmitter release. Single s-units were inserted between interneurons and motor neurons, and between pairs of motor neurons. Preliminary simulations showed that the time course of motor neuron synaptic potentials, which exhibit a fast and slow exponential decay, could not be produced by a local bending model having only one s-unit with a single time constant at each connection. Therefore, two s-units—one with a fast and one with a slow time constant—were inserted between sensory neurons and interneurons. This was consistent with intracellular recordings of synaptic potentials from interneurons in response to P cell stimulation, which showed both fast, slow, and mixed fast and slow synaptic potentials (Lockery and Kristan 1990b).

In equation 5.3, the sigmoidal function $f(V_i)$ represented the many processes which limit synaptic current, including limits on presynaptic calcium accumulation and number of vesicles released, the saturation of binding of transmitter to receptors, and the reversal potential of the synaptic conductance. The form of the sigmoidal function was determined by extrapolating from existing data on steady-state postsynaptic voltage as a function of presynaptic voltage at the synapses between two different inhibitory-excitatory motor neuron pairs (Granzow et al. 1985). Threshold for synaptic transmission was placed at the resting potential ($V = 0$), consistent with the fact that the inhibitory motor neurons do not release transmitter at rest (Lockery and Kristan, unpublished). Total current due to chemical synapses I_j^{chem} was given by:

$$I_j^{chem} = \sum_i w_{ji} S_{ji} \quad (5.4)$$

where w_{ji} is the strength of the chemical synapse to neuron j from neuron i , and is functionally similar to the product of the maximum synaptic conductance and the driving force ($E_{rev} - V_m$). Thus, in the model, currents from chemical synapses added linearly. Combining equations 5.1b through 5.4 yields the final equation:

$$\tau_j \frac{dV_j}{dt} = -V_j + R_j \left[\sum_i g_{ji} (V_i - V_j) + \sum_i w_{ji} S_{ji} \right]. \quad (5.5)$$

Input Resistance and Time Constants

Motor neuron input resistance (20 M Ω) and time constant (10 ms) were estimated from the time course and amplitude of the postsynaptic potential in DE in response to depolarizing current step (1.5 nA) injected into DI (Granzow et al. 1985). Interneuron input resistance was assumed to be

higher (40 M Ω) because the somata of most local bending interneurons are smaller than the somata of the motor neurons. For simplicity, the time constants for the interneurons were the same as for the motor neurons. The time constants for the s-units between motor neurons were chosen to reproduce the time course of postsynaptic voltage in DE upon injection of a square step of depolarizing current in the presynaptic motor neuron cell DI. For simplicity, the same value was chosen for the s-units for the connection from interneurons to motor neurons. This connection was chosen as the standard synapse because it is an established monosynaptic connection (Granzow et al. 1985); the same time constant was used for s-units between interneurons and motor neurons. Time constants for the fast (10 ms) and slow (1500 ms) s-units between sensory and interneurons were fit by hand, using a simplified model network comprising one sensory neuron, one interneuron, and one motor neuron. The values selected provided the best fit to the time course of the intracellularly recorded motor neuron synaptic potential in response to P cell stimulation.

Fixed Weights

The coupling resistance between pairs of electrically coupled motor neurons was calculated to achieve an effective coupling ratio of 0.1, in accordance with pairwise intracellular recordings from motor neurons (Friesen, unpublished; Lockery and Kristan, unpublished). Weights for chemical synapses between motor neurons were adjusted by hand to reproduce the steady-state response of the postsynaptic motor neuron to maximum depolarization of the presynaptic neuron (Granzow et al. 1985).

Optimized Weights

Connection strengths to and from interneurons were optimized by backpropagation to match the time course and amplitude of motor neuron synaptic potentials in the model to those in physiological recordings. The synaptic potential time course of a representative motor neuron recording was adopted as a template (figure 5.8B, upper traces). To construct the training set for backpropagation, the amplitude of the template was scaled to the average peak amplitude for each motor neuron in response to each of the eight patterns of single and paired P cell stimulation (Lockery and Kristan 1990a; figures 5.4 and 5.5). In accordance with the backpropagation algorithm, connections from sensory neurons to interneurons and interneurons to motor neurons were initially small and randomly assigned. For each iteration of the optimization procedure, an input pattern from the data set was presented and the network's behavior was simulated for 100 time steps (10 ms/step). The time course and amplitude of the simulated synaptic potential in each motor neuron was then compared to the desired time course for

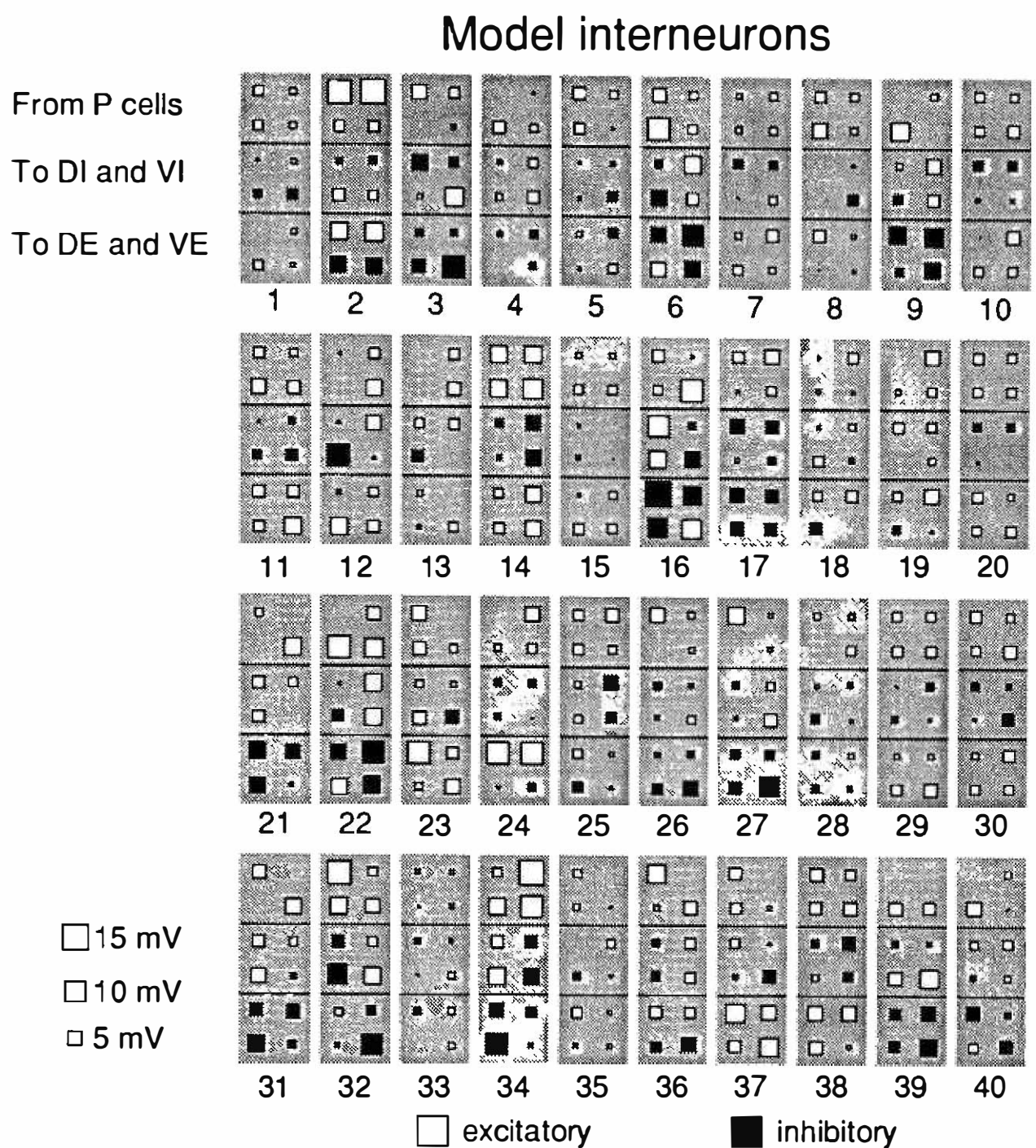
the relevant motor neuron in response to the same pattern of P cell stimulation; this gave a measure of the instantaneous error of the network at each time step. This procedure was repeated for each pattern in the data set. Each connection in the network was then adjusted by a small amount as specified by the algorithm and the process was repeated until the average instantaneous error was less than 0.19 mV.

Weights from sensory neurons to interneurons were constrained to be positive, since no inhibitory connections have been observed between such pairs. No constraints were placed on connections from interneurons to motor neurons. After 10 to 20 thousand iterations, a good match was obtained between simulated and desired motor neuron synaptic potentials (figure 5.8B, lower traces) for each input-output relation in the data set.

Performance of the Model

To compare interneurons in the model network to actual interneurons, simulated physiological experiments were performed. Interneuron input connections were determined by recording the amplitude of the postsynaptic potential in a model interneuron while each of the P cells was stimulated with a standard current pulse. Output connections were determined by recording the amplitude of the postsynaptic potential in each motor neuron when an interneuron was stimulated with a standard current pulse. This analysis yielded strong similarities between the model and biological network. First, model interneurons (figure 5.9), like those in the real network (figure 5.7B), generally received three or four substantial connections from P cells and thus received both dorsal and ventral P cell inputs. This means that interneurons in the model contributed to dorsal, ventral, and lateral local bends, as the distributed hypothesis holds. Second, individual interneurons had significant effects on seven or eight of the motor neurons. As in the biological network, the effect on an inhibitory motor neuron was not always the opposite of the effect on the excitatory motor neuron with the same projective field. For example, interneuron 12 (figure 5.9) excites both the right-hand DE and the right-hand DI. Finally, the range of fast, slow, and mixed synaptic potentials recorded in the interneurons in response to P cell stimulation was also reproduced by the model (figure 5.10). An interneuron in the model responded with a fast, slow, or mixed synaptic potential depending on the relative weights assigned to the fast and slow synaptic processes represented by the two different connections between each sensory neuron and interneuron in the model (figure 5.8A).

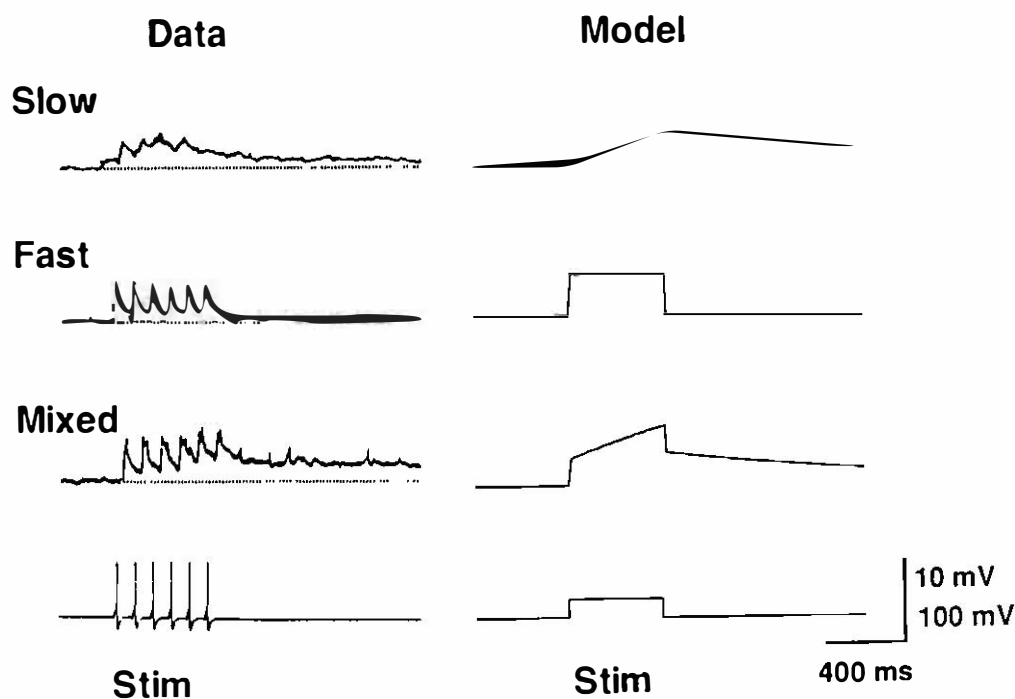
The qualitative similarity between the connectivity and temporal response properties of model and actual interneurons established the central result that a population of interneurons resembling the identified dorsal local bending interneurons could mediate local bending in a distributed



Input and Output Connections of Model Local Bending Interneurons

Figure 5.9 Model interneurons, like the actual interneurons, received substantial inputs from three or four sensory neurons and had significant effects on most of the motor neurons. Symbols as in figure 5.7.

processing system without additional interneurons specific for different forms of local bending. This constitutes a plausibility claim or an existence proof (Crick 1989) that the actual local bending circuitry could operate as we suppose. This was an important result because beforehand we thought that dedicated interneurons might be necessary to achieve accurate amplitudes and time courses for each motor neuron in response to each pattern of sensory input. Note that a negative result—failure of the optimization procedure to find a solution to the local bending input-output function—would have told us nothing. This is because the true absence of a solution cannot be distinguished from the case in which the solution is simply missed by the optimization procedure. This is a general problem with optimization via gradient descent, since it can become trapped in a local minimum, i.e., stuck with a set



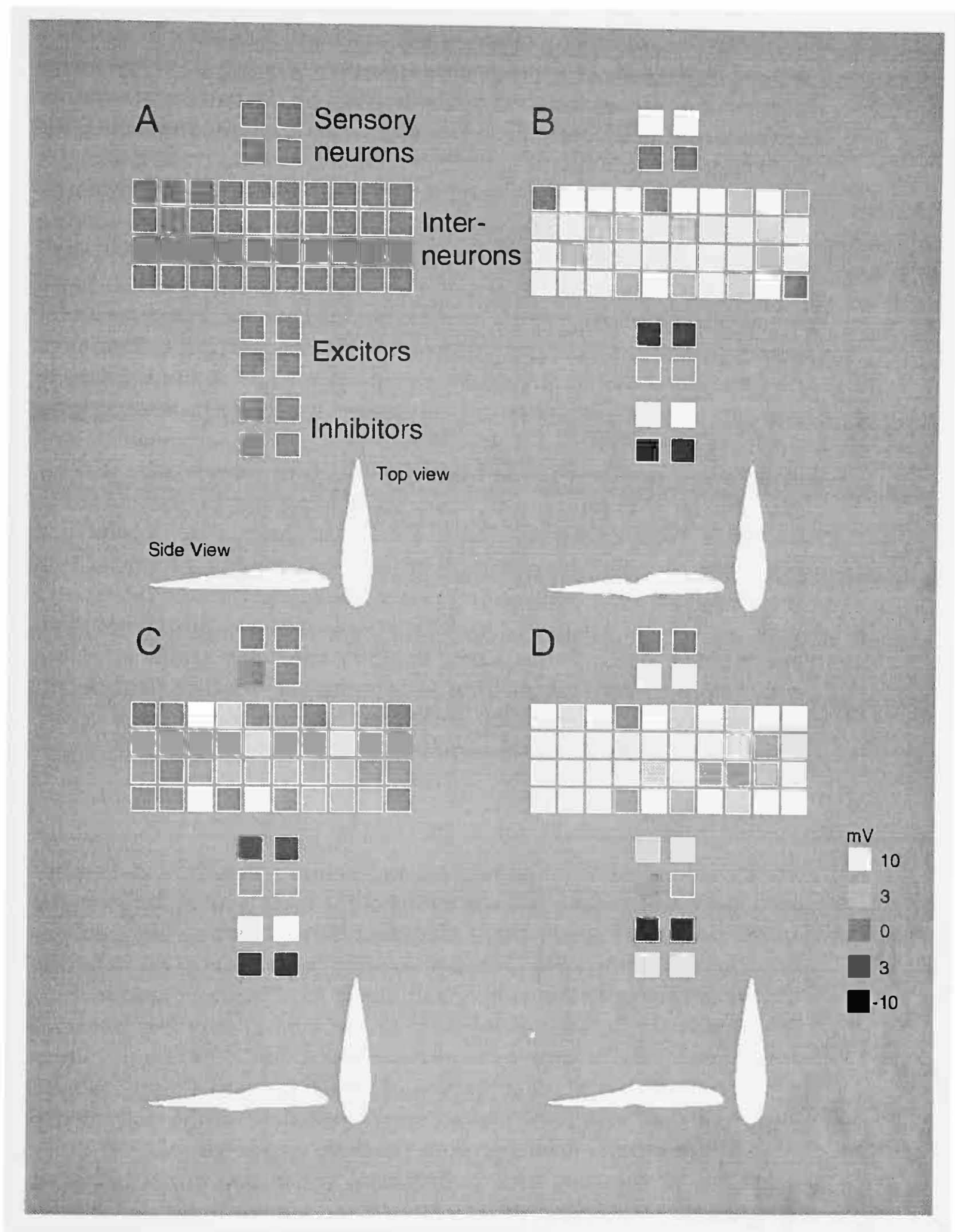
Recorded and Simulated Synaptic Potentials

Figure 5.10 Comparison of recorded and simulated synaptic potentials in interneurons in response to stimulation (Stim) of a single P cell. (Data) Physiological recordings are shown on the left. (Model) Simulated synaptic potentials are shown on the right. Interneurons with fast, slow, and mixed responses were seen in physiological recordings; similar types were seen in the model after training. In the simulations, the trains of P cell impulses were modeled as a step increase in average membrane potential in the sensory neuron. (Reprinted from Lockery and Sejnowski 1990 with permission.)

of connection strengths that are only a partial solution to the problem. The danger of local minima can be mitigated to some degree by training many networks from different random initial weight values, and by using different algorithms (Kirkpatrick et al. 1983).

The construction of a complete and realistic model of the local bending reflex makes it possible for the first time to visualize the response of the entire network as it evolves in time (figure 5.11). This analysis showed that in the model almost every interneuron contributed to the rapid rising phase of the synaptic potentials in the motor neurons (figure 5.11B). By contrast, the slow falling phase was under the control of many fewer interneurons (figure 5.11C). This observation suggests that, in the model, a single pattern of sensory input is represented by different patterns of interneuron activation at different times relative to stimulus onset and offset. Similarly, a single pattern of motor neuron excitation and inhibition is represented by different interneuron activity patterns at different times relative to the stimulus. The concept of time-dependent sensory and motor representations has received relatively little attention because models of sensorimotor integration have tended to focus on static representations. The idea of time-dependent representations can now be tested by physiological experiments to determine whether the biological network exhibits a similar time dependence.

Comparing the response to different patterns of P cell stimulation showed that most interneurons contributed to more than one form of



The Response of the Complete Local Bending Network as a Function of Time

Figure 5.11 (A) Resting. (B) Rising phase of the response to dorsal stimulation. (C) Falling phase of the response to dorsal stimulation. (D) Rising phase of the response to lateral stimulation. In each panel, the shading of the squares shows the voltage of the indicated sensory, motor, or interneuron, according to the gray scale at the left. The leech's profile is shown in side view, with anterior to the left, and in top view, in which anterior is up.

local bending, but each pattern of sensory input was associated with a unique spatial pattern of interneuron activation (figure 5.11D). The spatial pattern of interneuron activation is decoded by the pattern of synaptic connections to the motor neuron such that each interneuron activation pattern generates a unique—and behaviorally correct—pattern of motor neuron excitation and inhibition. An interesting theoretical question is whether distributed representations across as many as 40 interneurons are necessary. Two possibilities arise. First, the large number of interneurons may be required to accommodate additional complexities in the input-output function of the reflex. So far, we have measured the motor output in response to eight patterns of single or paired P cell stimulation, and these measurements were made at a single stimulus intensity and duration. Clearly, there are many more input patterns to investigate as one varies the number of P cells and the stimulus parameters. These experiments may reveal additional input-output complexities that explain the large numbers of interneurons that appear to be involved in the reflex. Second, each interneuron may also participate in behaviors other than local bending. Indeed, cell 115 was initially identified as a member of the swimming central pattern generator (Friesen 1989a), and both cell 115 and 125 have detectable effects on shortening (Wittenberg and Kristan 1992). Participation in other behaviors could force the connections of each interneuron to compromise between its different behavioral roles, necessitating additional interneurons to make up for the attendant functional shortcomings.

CONCLUSION

We have combined compartmental and neural network modeling approaches to construct a realistic model of the local bending reflex in the leech. The model showed that identified dorsal bending interneurons are consistent with a distributed processing hypothesis of the network. The model suggests that subpopulations of local bending interneurons that remain to be identified—such as the ventral bending interneurons—are similar in kind to the interneurons identified so far. Addition of more physiological constraints would make the model more realistic, and could lead to specific and accurate predictions of the connectivity of as yet unidentified interneurons (Lockery and Sejnowski 1992). For example, we can now train networks in which one group of model interneurons is constrained to have all the known input and output connections of the dorsal bending interneurons. A second, undifferentiated group of model interneurons would then be free during training to develop connections that predict the complementary input and outputs of the actual interneurons that act in concert with the identified dorsal bending interneurons. These predictions could then be tested by identifying the remaining local bending interneurons in the ganglion.

Backpropagation and other network training algorithms enable us to

construct working models of distributed processing networks that were not possible a few years ago. By using neural network units to represent electrical and synaptic compartments, it is possible to make these models as realistic as compartmental models themselves. In the future, we can expect this approach to be applied to other issues in systems neuroscience, including mechanisms of central pattern generation in invertebrates and vertebrates, the processing of temporal information, and the prediction of sites of synaptic plasticity. This chapter has focused on networks of neurons. However, the network modeling approach has potential biological applications far wider than the nervous system, since the units in the network can represent other biological variables, including environmental (Lockery and Sejnowski 1991) and developmental factors (Mjolsness et al. 1991), as well as aspects of the immune response (Vertosick and Kelly 1991).

ACKNOWLEDGMENT

Supported by a Postdoctoral Fellowship from the National Institutes of Health, The Howard Hughes Medical Institute, The Bank of America-Giannini Foundation for Medical Research, and The Mathers Foundation.

6

Neural and Peripheral Dynamics as Determinants of Patterned Motor Behavior

Hillel J. Chiel and Randall D. Beer

Patterned motor behavior is ubiquitous among animals. Examples of patterned behaviors include chewing, walking, and escaping from predators. Equally striking, however, is that these patterns of behavior do not recur rigidly. Instead, they vary in response to moment-to-moment changes in an animal's body and environment, in such a way that the animal is usually able to continue to carry out that behavior. For example, animals are capable of feeding steadily, even though the properties of food continually change as it is consumed. Changes in the slope of the ground, or small obstacles, do not usually disrupt locomotion. Even simpler organisms show this flexibility. For example, cockroaches can make appropriate escape responses if they encounter obstacles during their escape turn (Ritzmann et al. 1991).

Two questions naturally emerge from these observations: First, what are the biological mechanisms allowing animals to generate patterned motor behavior that can be rapidly adjusted to continuously changing external and internal conditions? Second, what useful principles could be derived from biological organisms for the purpose of constructing devices with these properties? This chapter will address both of these questions.

Attempts to answer these questions have focused primarily on the properties of the nervous system. We suggest that a broader perspective is necessary. Nervous systems and the bodies in which they are embedded are each discrete entities with distinct properties. At the same time, the two are in continuous mutual interaction, and it is principally these interactions that are responsible for an organism's behavior. From one perspective, the nervous system is continuously bombarded by sensory feedback both from the body itself and from the environment, whose inputs are filtered and transformed by the body. From another perspective, the body is continuously receiving output from the nervous system, which changes the arrangement of its parts, in turn leading to new sensory inputs to the nervous system. Evolution has selected for the appropriate relations between these entities so that, working in conjunction, they generate behavior that enhances the survival and reproduction of an organism. As we hope to demonstrate, this broader perspective

has important implications for the understanding of biological nervous systems, as well as for the design of artificial neural networks and of autonomous, adaptive agents.

To support this viewpoint, the next three sections of this chapter will review the evidence that the properties of nervous systems, of the peripheries they control, and their interactions are essential for the generation of adaptive patterned behavior. We begin with a review of the complexity of biological neural networks. Neurons are themselves complex dynamic systems, due to their intrinsic properties, and they are connected together in highly ordered and complex architectures, which further enriches their repertoire of dynamics. We next review the complexity of the dynamics of different peripheral plants. As is conventional in engineering terminology, we use the term *plant* herein to include the elements of the periphery responsible for generating movements and forces. These elements are highly nonlinear and are also interconnected in complex architectures. We then review the adaptive interactions that can occur between the nervous system and its periphery during behavior, and focus on a specific example, describing a biologically inspired neural network for controlling hexapod locomotion. This example is used to illustrate the utility of this perspective for understanding both biological pattern generators and engineering controllers for robot locomotion. The chapter concludes with a general discussion of the implications of this perspective both for improving our understanding of the biological mechanisms of adaptive behavior and for deriving general principles for the design of adaptive devices.

NEURAL NETWORK DYNAMICS

Complexity of Biological Nervous Systems

What are the essential properties of nervous systems that need to be understood in order to explain patterned, adaptive behavior? Twenty years ago, many neurobiologists felt that the “wiring diagram” for the circuit that controlled a behavior of interest was all that was necessary to understand it. That is, if the sensory neurons, motor neurons, and interneurons controlling a behavior were identified and the synaptic interactions between them had been determined, the circuit was completely defined. Furthermore, it was assumed that the essential information communicated between neurons was carried by action potentials. Finally, it was assumed that information in neural circuits was primarily processed in a linear fashion, through a series of “stages” or “stations.” Given these assumptions, the architectures that were proposed for the Perceptron (Rosenblatt 1958) or for multiple layer feedforward neural networks (Rumelhart and McClelland 1986) were reasonable models of the nervous system.

In the last two decades, however, our understanding of the function

of the nervous system has undergone significant changes that have fundamentally altered this view. First, it has become evident that neurons (both in vertebrates and invertebrates) possess rich and complex intrinsic properties. Most neurons have multiple channels to different ionic species, and these channels can be regulated in a wide variety of ways: They can be turned on or off by voltage, molecules, or ions. Some of these channels can be active in the absence of external inputs to the cell, and endow it with a variety of dynamic properties, such as the ability to oscillate (Llinás 1988; Selverston 1988; Yamada et al. 1989). Thus, it is not enough to specify the inputs to a neuron to predict its outputs; its internal state will also determine its behavior. As a consequence, neurons may be better represented as nonlinear dynamic systems in their own right. For example, the intrinsic conductances of thalamic neurons can allow them to act as linear input/output devices, relaying information directly to cortex, but when they are hyperpolarized, these conductances cause the neurons to burst, significantly transforming their inputs (Llinás 1988). In terms of the model neurons that have often been used in artificial neural networks, the input/output relationship would need to be represented as a function both of voltage and of time.

Second, the interactions between neurons are complex. The differential distribution of synapses on complex dendritic trees of neurons can significantly affect the nature and intensity of their inputs to a neuron. In addition, synapses may have multiple time courses (e.g., initial excitation, slower inhibition, and still slower excitation [Getting and Dekin 1985]), and connections may be dynamically reconfigured (e.g., by inhibition of specific neurons [Getting and Dekin 1985], or by the actions of neuromodulators [Harris-Warrick and Marder 1991; Marder and Hooper 1985]). Receptors controlling the synaptic response may be gated both by the presence of a chemical, such as a neurotransmitter, and by voltage, so that the synaptic connections between neurons can be affected both by their own activity and by the activity of neurons impinging on them (discussions of these and other complexities in synaptic interactions are found in chapters 2, 3, and 4). Influences may occur over a variety of spatial and temporal scales: A neuromodulator which is only slowly broken down may affect a very large number of neurons in its vicinity over an appreciable period of time as it diffuses away from its point of release. Furthermore, such compounds are likely to selectively activate those subgroups of neurons that have a receptor for that substance. Neuromodulators may also have subtle but profound effects on the intrinsic properties of neurons, activating or inactivating intrinsic currents and thus changing their “electrical personality.” Field potentials may alter the excitability of neurons in different regions of the brain (Nunez 1981). To recast this in terms of the model neurons that have been frequently used in artificial neural networks, the input/output relationship, threshold for activation, bias currents, and connection weights

may all be time-varying in complex but predictable ways. Thus the “functional wiring diagram” for the circuit may be systematically changing over time as the circuit operates, even in the absence of learning.

Third, it has become clear that many neurons communicate with one another without generating action potentials, releasing transmitter in direct proportion to the extent of their depolarization (Burrows 1989). Electrotonic decrement of membrane potential in a complex dendritic tree makes it unlikely that synaptic inputs will be equally communicated to all parts of a neuron. Thus, the shape of a neuron may play an important role in transforming inputs to outputs, along different parallel pathways.

Fourth, it has become clear that feedback connections are nearly ubiquitous in neural circuits. Many sensory modalities receive feedback from central neural structures. Motor systems possess complex feedback connections. Many central neural structures have extensive intrinsic feedback. For example, the bulk of cortical connections are feedback connections within the cortex itself (Shepherd 1990a). Feedback may also occur through nonsynaptic mechanisms. For example, field potentials generated by surrounding nerve cells may change the excitability of a neuron, whose activity can in turn enhance or depress the population field potential. Thus, a serial information processing model of the nervous system is likely to be a significant distortion of the actual mechanisms of processing in the brain. Other aspects of feedback and recurrent networks are discussed in chapters 5 and 7.

Feedback connections illustrate the highly organized architectures of biological neural networks. Many artificial neural networks currently in use begin with the assumption that interconnections are randomly weighted and then tuned by a “learning rule” or “training algorithm.” Far from being randomly interconnected at the time an organism begins to function, nervous systems are highly organized by developmental processes, and much of their connectivity is specified by the time that an animal is born. Insects, for example, are capable of walking almost immediately after hatching. Although it is undoubtedly the case that experience shapes and selects preexisting connections, and that new connections may form during the lifetime of an organism, this modification of connectivity occurs in the context of a complex preexisting structure.

Many of these points have been eloquently summarized by others. Bullock (1980) proposed the notion of the “neural throng” as a descriptor of the richness of neuronal interactions, as opposed to the notion of a “neural circuit”, emphasizing the multiplicity of complex interactions that actually occur between nerve cells (Bullock 1980). More recently, Getting and Dekin (1985) proposed the notion of the “polymorphic network”, capable of generating a multiplicity of functionally distinct neural circuits in different behavioral contexts. They based this concept on their experimental studies of the neural control of escape swimming in *Tritonia*, which indicated that the same neurons could function to

generate alternating dorsal and ventral flexions during an escape swim, while generating simultaneous dorsal and ventral contractions during a withdrawal response. Further support for this viewpoint has come from recent work on the stomatogastric ganglion of the crab *Cancer borealis*. Motor neurons that are part of this ganglion not only innervate multiple parts of the stomach musculature, but their activity is dynamically reconfigured so that they participate in a wide variety of motor patterns (Weimann et al. 1991).

Summary: Network Dynamics

This brief review of our current understanding of biological neural networks suggests not only that neurons and nervous systems are complex, but that this complexity is essential for their function. Neurons have complex intrinsic properties, multi-action synaptic connections that can act over a variety of temporal and spatial scales, and graded interactions that depend on their shapes, and are organized by evolution and by developmental processes into complex architectures. These properties give rise to the extremely complex nonlinear dynamics of biological nervous systems, which may not be adequately captured by artificial neural networks that ignore these features. An example of the rich dynamics that can emerge when some of these properties are incorporated into an artificial neural network is described later in this chapter.

PERIPHERAL DYNAMICS

Complexity of Peripheries

Fish and eels swim, horses and centipedes walk, birds and insects fly. Animals locomote in a variety of different ways, using strikingly different peripheries. The differences in body design may reflect the ecological niche that an animal inhabits or the evolutionary history of that species. It is hardly surprising, therefore, that pattern generators manifest a wide variety of architectures, even when the end result is alternating contractions of different body parts for the purpose of locomotion (Kristan 1980). Furthermore, the same periphery may be utilized to execute qualitatively different behaviors. For example, a variety of animals, such as turtles and cats, utilize the same limb musculature for locomotion and scratching (Stein 1983). As a consequence, it is not possible, in general, to understand the architecture of a biological neural network without understanding the periphery it controls, nor is it possible to predict the behavioral capabilities of a given periphery without understanding its neural control.

The peripheral plant may also create significant constraints on possible solutions that a nervous system has at its disposal for a given problem. For example, Carrier (1991) has shown that, because the hypaxial

muscles need to work in opposition during running and in synchrony during breathing, lizards cannot breathe and run at the same time.

These observations suggest that the design of a neural architecture is intimately tied to the periphery it controls, and that the search for general neurobiological circuit design principles will fall short if the periphery is ignored. They also suggest that a given neural network will act very differently if it does not receive the sensory and proprioceptive feedback signals that shape its activity in the intact animal. For example, Pearson and his colleagues showed that the duration and frequency of locust flight were significantly altered in the absence of feedback from wing stretch receptors during the appropriate phase of movement (Pearson et al. 1983). Indeed, given that neural controllers have co-evolved with the body that they control, it is likely that many aspects of biological neural networks only make sense in the light of the properties of the periphery.

What are the properties of the peripheries of animals that may be relevant to neural control? Biologists have identified three broad categories of solutions for moving and generating forces (Kier and Smith 1985; Wainwright et al. 1976): (1) hydrostatic skeletons, such as the body of an earthworm, (2) muscular hydrostats, such as the tongue, and (3) musculoskeletal systems, such as the human arm. Combinations of these solutions are found in most animals. For example, during feeding, we utilize our jaw musculature (part of the musculoskeletal system) for crushing and grinding food, at the same time that we use our tongue (a muscular hydrostat) to reposition food within the oral cavity. What is common to these different solutions is the force-generating material, which is muscle. We will briefly discuss the properties of muscle before turning to a more detailed discussion of these different peripheral plants.

Muscle

Muscle has springlike properties. Even in the absence of neural inputs, when muscle is stretched, it will tend to resist that stretch by generating a restoring force. This passive stiffness of muscle allows it to resist rapid, large perturbing forces. When muscle is activated by neural inputs, its stiffness can be significantly modulated by the frequency of inputs to each of its fibers and the number of fibers that are recruited (Ghez 1991).

Muscle properties vary (Shepherd 1988). Skeletal muscle can generate its own action potentials, thus uniformly depolarizing the entire muscle fiber, and requires repeated neural activation to maintain contractions. It can develop high forces rapidly (at a high energetic cost), or smaller forces more slowly (at a lower energetic cost). Smooth muscle (and much invertebrate muscle) does not generate its own action potentials, so that different regions can be differentially activated. Although smooth muscle generally contracts slowly, some smooth muscle can go into a “catch” state in which molecular cross-links are formed within the muscle

fiber, and no further neural stimulation or energetic expenditure is required to maintain the contracted state. Clams, for example, use the “catch” state to keep their shells firmly closed against predators and waves (Wilson 1979). Thus, the “actuators” used by biological organisms have highly non-linear properties that are exploited during behavior.

Hydrostatic Skeletons

Many organisms utilize structures surrounded by muscle and filled with fluid to move and generate force. Many simple wormlike organisms utilize these structures, but examples are found in higher organisms as well (Chapman 1958, 1975).

Hydrostatic skeletons serve as a particularly effective example of the importance of the dynamics of the periphery. In general, their volume cannot change, and their internal pressure must be constant throughout. Thus, in order to expand one part of the body, another part must be contracted. These properties make it difficult to use them to exert point forces at particular locations, as can be done effectively by musculo-skeletal systems. However, even in the absence of neural inputs, these structures will assume and maintain particular shapes. Thus, the neural control of such structures may exploit this tendency or compensate for it, depending on what behavior the animal is doing.

Wadepuhl and Beyn (1989) have recently published a quantitative model of a hydrostatic skeleton that is designed to serve as a model of leechlike organisms. The model organism consists of repeated hexahedral elements. Each edge of the hexahedral element is contractile, modeling the circumferential and longitudinal muscles found in the leech body wall. The authors demonstrate that lengthening the structure requires constriction of both longitudinal and circumferential elements, and that the structure can assume curved shapes by differential contractions of longitudinal muscles. Because changes in volume in one part of the structure propagate throughout it, the effects of contracting different regions can be very nonintuitive.

Another study of hydrostatic skeletons has recently been published by Erdos and Niebur (1990), who modeled the biomechanics and neural control of locomotion in the nematode *C. elegans*. The forces encountered by the animal are quite different depending on the surrounding water level. If the water level is high, the animal is subjected to relatively low surface tension, and swims. If the water level is low, but above the surface of the animal’s body, the animal slips. Finally, if the water level is low enough, the animal’s body creates a groove in the water due to surface tension, and the animal creeps. Thus, one implication of these observations is that the movements of the periphery may give rise to somewhat different behaviors depending on environmental conditions.

Understanding the biomechanics of hydrostatic skeletons can aid in

the analysis of animal behavior. For example, Teyke and co-workers (Teyke et al. 1990) studied the mechanisms of part of the appetitive response of *Aplysia* to food—orienting its head to a piece of seaweed. The response of an animal to a brief contact with food could be modeled well by representing its head and neck as a fluid-filled cavity subjected to a brief impulse. Indeed, based on this model, they were able to demonstrate that the accuracy of the animal's response to food under closed-loop conditions (i.e., when the animal maintained contact with the food as it responded to it) was probably due to response substitution. That is, an initially large movement in response to food at the edge of the lips is replaced by progressively smaller movements as the food contacted the inner parts of the lips during the movement. Thus, understanding the properties of the periphery led to interesting suggestions about the neural control of the behavior.

Muscular Hydrostats

More complex animals possess flexible structures that are made solely of muscle. Examples of such structures would be the human tongue, the tentacle of an octopus, or the trunk of an elephant. Since the volume of muscle changes very little during contraction, such structures have been termed muscular hydrostats (Kier and Smith 1985).

Different arrangements of muscle fibers are key to the ability of muscular hydrostatic structures to flexibly move throughout three-dimensional space, to conform to irregularly shaped objects, and to insert themselves into tortuously shaped spaces that would be inaccessible to jointed devices. In general, these structures contain muscle fibers with very specific geometric arrangements. Thus, in some tongues, circumferential fibers surround longitudinally oriented fiber bundles. In others, transverse and horizontally oriented muscle fibers surround bundles of longitudinally oriented fibers. In other muscular hydrostatic structures, such as the elephant's trunk, some muscle fibers wind helically around the length of the structure.

Understanding the biomechanics of muscular hydrostats has also provided insights into animal behavior. For example, Wilson and co-workers (Wilson et al. 1991) have used a continuum model of a muscular hydrostat to analyze the properties of an elephant's trunk during lifting. They were able to calculate the stiffness of the elephant's trunk as it lifted a known weight. An interesting observation they made was that the elephant prestresses its trunk during lifting; that is, it grasps the load and shortens before it begins to lift. They hypothesized that this increased the forces it can exert during lifting.

Chiel and co-workers (Chiel et al. 1992b) have recently developed a model of another muscular hydrostat, a reptilian tongue, and used the model to study lapping behavior in the lizard *Tupinambis nigropunctatus*. The model tongue consisted of two muscles: a longitudinal muscle

whose contractions shortened the tongue, and a circumferential muscle wrapped around the longitudinal muscle whose contractions squeezed the longitudinal muscle and lengthened the tongue. The cylindrical geometry of the structure, its constant volume under deformation, and a pressure vessel model of the hydrostat were used to relate lengths and forces in the longitudinal muscle to those in the circumferential muscle. Using anatomical parameters derived from measurements of the lizard's tongue, electrophysiological data measured from the lizard, and biomechanical data from cat hindlimb muscle (since none was available for the lizard), they studied the static and dynamic properties of the structure. Because the biomechanical data were obtained from cat muscle, they examined the effect on the model of varying critical parameters, and found that the model was robust to such variations.

When the forces exerted by the circumferential muscle were converted to forces exerted in the longitudinal direction, the model revealed an unexpected result. As the tongue lengthens, the force exerted by the circumferential muscle falls precipitously. As a result, the final length that the tongue can achieve will depend on the load it is attempting to manipulate. Indeed, cineradiographic studies of the lizard's tongue (Smith 1984) demonstrated that it lengthened greatly when it was dealing with light loads (such as lapping water), but remained short when it was dealing with heavy loads (e.g., when the lizard was engaged in pharyngeal packing). Furthermore, when it is long, the tongue cannot be very stiff. Finally, the model indicated that the form of the passive stiffness could significantly affect the tongue's behavior. These properties of the structure are likely to have a profound effect on its neural control.

In the last few years, Wilson and co-workers have begun to develop continuum models of deformable structures, such as bellows, pneumatically actuated devices, and solid structures made of highly elastic material, because of their potential for applications to robotics. They believe that these structures could be used as actuators for robots, and have very different capabilities than jointed manipulators, such as the ability to conform to irregular objects or reach into areas that might be inaccessible to jointed manipulators (Wilson 1984). Studying the neural control of these structures in biological organisms may provide useful ideas for the design of controllers for artificial muscular hydrostats.

Musculoskeletal Systems

The most familiar peripheral plant is the musculoskeletal one, underlying the construction of much of the human body, as well as many other complex organisms. The control problems associated with these peripheries are no less complicated than those associated with hydrostatic skeletons or muscular hydrostats. Muscles span the joints between different bones, in some cases spanning a single joint, in others spanning multiple joints. Generally, muscles are organized in agonist/antagonist pairs and

can be activated independently, or co-contracted. Independent activation allows torques to be generated around specific joints; co-contraction can act to increase the stiffness of a joint.

If the angles and torques at each joint of a multijointed limb are known, it is straightforward to compute the position and the force at the endpoint of the limb. In general, however, the limbs of most animals are capable of reaching a given position (or generating a given force) in a multiplicity of ways, because the joints making up the limb have excess degrees of freedom. To move a limb in a three-dimensional space requires at least 3 degrees of freedom; but, to take a specific example, the human arm has 7 degrees of freedom (3 at the shoulder, 2 at the elbow, and 2 at the wrist). Thus, to specify a position and force for the endpoint of the limb by determining the angles and torques of the constituent joints (the inverse kinematics and inverse dynamics problems) is an ill-posed problem, because it does not have a unique solution. Both motor physiologists and roboticists working with jointed manipulators have increasingly realized the importance of dealing with this aspect of the periphery. In the next section of this chapter, we will discuss some of the ways in which nervous systems appear to cope with these problems.

Finally, it is important to point out that no set of muscles ever works in isolation. Other muscles must be utilized to provide a platform against which forces or movements can be generated. Indeed, the effects of activating a set of muscles can be radically different depending on the conditions of these postural muscles. We are all familiar with the consequences of attempting to throw a heavy object if we are insufficiently braced—we may lose our balance and fall backward. The nervous system must therefore deal with the control of these platform or postural muscles as it generates behavior.

Summary: Peripheral Dynamics

Understanding the properties of the periphery is crucial to understanding how nervous systems work because they have co-evolved. Nervous systems must control highly nonlinear actuators (muscles), which are organized in complex architectures that have their own dynamic properties. Examples of such peripheral structures are hydrostatic skeletons (e.g., those of worms), muscular hydrostats (e.g., tongues), and musculoskeletal systems (e.g., the human arm). Each of these different peripheral plants pose unique problems for control, as well as providing the nervous system unique opportunities for generating certain kinds of behavior with great ease.

NEURAL/PERIPHERAL INTERACTIONS

The two previous sections of this chapter have provided evidence that the complexity of biological nervous systems is essential for their ability

to generate robust, flexible patterned behavior, and that properties of the periphery may be crucial for understanding an animal's behavioral capabilities. In this section, we argue that it is the ongoing interactions between the nervous system and its periphery that are jointly responsible for adaptive behavior.

This broader viewpoint on the problem of patterned motor behavior has several important implications, which we will also discuss. First, it may not be possible to understand how a nervous system generates behavior unless one studies the periphery that it controls. Second, many properties of the nervous system can be better understood in the light of its relationship to the periphery.

Musculoskeletal Biomechanics and Spinal Cord Organization

The study of vertebrate motor control has provided experimental evidence for the validity of this perspective. Because of the agonist/antagonist arrangement of muscles around joints, when a limb is disturbed the mechanical stiffness of the musculature will generate a restoring force that will tend to return the limb to its initial position. Since the restoring forces at this point will be zero, it is an equilibrium point. Thus, one mechanism whereby nervous systems generate trajectories of limb movement could be by continuously changing the location of the equilibrium point, which can readily be done by changing muscle stiffness.

Experimental results suggest that interneurons in the spinal cord group motor neuronal activations to create these equilibrium points. Bizzi and co-workers have studied this problem using the scratch reflex of the frog (Bizzi et al. 1991). They have obtained experimental evidence that the spinal cord of frogs, even when isolated from the brain stem, is capable of generating convergent force fields when it is stimulated, i.e., force fields that induce a limb to move to a particular equilibrium position. They have also shown that these convergent force fields are not due to organization of motor neurons themselves or to random activation of different motor neuron pools, and are observed in the absence of afferent feedback. Furthermore, stimulation of two different points in the spinal cord generates a force field which is a sum of the vectors specifying each individual force field.

These data strongly support the hypothesis that the nervous system handles the excess degrees of freedom problem by specifying sequences of equilibrium points. Furthermore, they suggest a mechanism whereby "higher order" inputs can act to shape and modify ongoing activity in response to proprioceptive and sensory feedback, thus utilizing the excess degrees of freedom when necessary. As Bizzi and co-workers (Bizzi et al. 1991) conclude:

The vectorial combination of motor outputs derived from activation across different areas of a very coarse map is a remarkable mechanism for producing a vast repertoire of motor behavior in a simple fashion.

Thus, these data strongly support the hypothesis that ongoing interactions between the nervous system and its periphery are essential for the generation of behavior.

Periphery May Be Essential for Normal Function of the Nervous System

One implication of the view that the ongoing interactions between the nervous system and its periphery are jointly responsible for adaptive behavior is that the nervous system may not function normally in the absence of its periphery. A crucial step in advancing the understanding of the neural control of behavior has therefore been the development of reduced preparations which produce the behavior of interest, and which simultaneously allow the investigator to record from neurons of interest. For example, the study of the neural basis of escape swimming in *Tritonia* depended on the development of a semi-intact preparation in which it was possible to record from motor neurons as the animal performed the behavior. In contrast, research into the neural basis of feeding behavior in *Pleurobranchaea* was hampered by the absence of an appropriate reduced preparation (McClellan 1982a, b). The wide range of patterns generated by the stomatogastric ganglion is not observed when it is studied in isolation from the periphery. Much recent exciting progress in understanding the function of this ganglion has come from the analysis of its dynamic rewiring by neuromodulators in response to changes in sensory feedback from the periphery (Katz and Harris-Warrick 1990).

Peripheral Properties Clarify Properties of the Nervous System

Another implication of this viewpoint is that many properties of the nervous system only make sense in the light of peripheral properties. Our understanding of the motor neuron serves as a good example. Sherrington's "final common pathway" was often regarded as a simple "read out" device for the central nervous system, faithfully conveying central commands to the musculature. Studies of motor neuronal properties have revealed that they have complex intrinsic properties, such as plateau potentials. These properties aid them in adjusting their outputs to the periphery (Kiehn 1991).

Properties of sensory neurons may also make more sense in the light of their relationship to the periphery and the behavioral context in which they are utilized. For example, an identified histaminergic neuron (C2) in *Aplysia* has been shown to be a mechanosensory neuron with unusual properties that make sense in context. The neuron transduces mechanical stimuli applied to the lips of the animal, as well as mechanical deformations of the lips due to the animal's own movements. This neuron also can excite or inhibit motor neurons for the extrinsic muscles of the buccal mass, which appear to serve a postural role during feeding

behavior. In addition, this neuron can presynaptically inhibit outputs of the central motor program for feeding as it impinges on these motor neurons, as well as itself being phasically inhibited by this central motor program. One explanation for these properties is that the neuron serves to monitor local conditions at an animal's lips during feeding, and can rapidly alter the intensity of the animal's response during one phase of feeding as the toughness or texture of food changes, dynamically altering that part of the motor program to respond to changes in local conditions (Chiel et al. 1990).

Patterns of neural activation of effector muscles also make more sense in the light of the properties of the periphery. For example, Smith's electromyographic studies of the tongue of *Tupinambis nigropunctatus* showed that there was considerable overlap in the timing of activation of both the circumferential and longitudinal (hyoglossus) muscles during lapping movements (Smith 1984). What is the significance of this co-contraction? The model of the lizard's tongue (described in the previous section) suggested that the co-contraction of the two muscles played several important roles in lapping. First, the rapidity with which the tongue retracts is due to the co-activation of the longitudinal muscle while the circumferential muscle is still contracting. Second, the ability of the tongue to retract as far as it does is due to the residual activation of the circumferential muscle, which tends to enhance the ability of the longitudinal muscle to exert force by stretching it.

Summary: Neural/Peripheral Dynamics

The work that we have reviewed strongly supports the hypothesis that adaptive behavior emerges from the joint interactions of the periphery and the nervous system. Thus, the organization of the spinal cord may reflect the biomechanical properties of the musculoskeletal system that it controls. Other experimental results suggest that, in the absence of the periphery, the nervous system may not function properly. Finally, experimental evidence suggests that properties of motor neurons, sensory neurons, and activation patterns of groups of motor neurons may only make sense in the light of the properties of the periphery.

A BIOLOGICALLY INSPIRED CONTROLLER FOR HEXAPOD LOCOMOTION

To illustrate the approach we have taken in this chapter, we will describe in some detail a neural network controller based on neurobiological data which we have used to control hexapod locomotion in a simulated insect and in a robot. We will begin by describing a heterogeneous neural network for locomotion, which is inspired by neurobiological data on insect walking; we will then describe the periphery in which it is embedded, and finally describe the interactions between the central

controller and its periphery. The locomotion controller provides another illustration of the importance of the ongoing interactions between the nervous system and its periphery in the generation of patterned motor behavior.

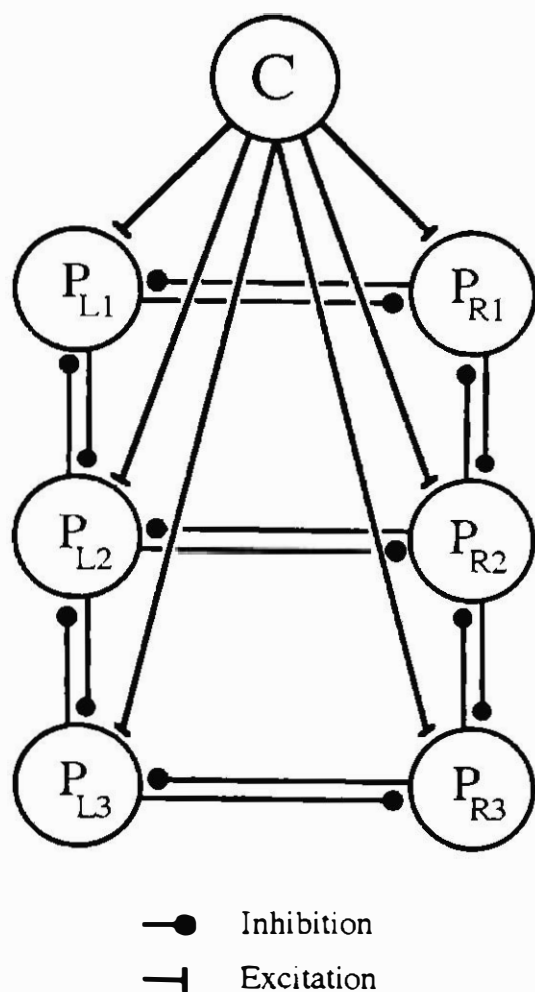
Neural Network Controller for Hexapod Locomotion

In studies of cockroach locomotion, Pearson and his colleagues obtained evidence that the protraction of the legs, i.e., the time during which they swung forward, was probably due to the activity of nonspiking oscillatory neurons (Pearson et al. 1973). Their studies did not indicate whether this oscillator was a single neuron or a coupled network of neurons, but the lack of overlap between times of activation of adjacent legs suggested that inhibitory connections existed between oscillators for adjacent legs.

Based on these results, we proposed a neural network architecture for the control of locomotion (Beer et al. 1989). Since the oscillatory system had not been specified experimentally, we created a model pacemaker neuron. Two intrinsic currents controlled the behavior of the pacemaker: One induced it to remain depolarized for a fixed period of time, and a second induced it to remain hyperpolarized. The duration of the second intrinsic current varied as a function of the voltage of the pacemaker, so that when the pacemaker was hyperpolarized, the time between bursts was lengthened, and when it was depolarized, the time between bursts was shortened. In addition, brief excitatory inputs during the time a pacemaker was off could turn it on and reset the phase of its bursting; brief inhibitory inputs when it was on could similarly reset its bursting phase. These pacemaker properties have been observed experimentally in nerve cells (Kandel 1976).

The neural network consisted of six pacemakers, one for each leg. Interactions between pacemakers were modeled after the interactions that had been deduced to exist between the flexor burst generators: Adjacent pacemakers inhibited one another. Finally, as a model of higher-order influences that are known to activate locomotion, we connected a single command neuron to all six pacemakers. Its steady excitation or inhibition of the pacemakers caused them to oscillate at different frequencies, due to their intrinsic properties (figure 6.1).

If the command neuron provided identical inputs to each pacemaker, we found that only one pattern of activity was generated by the network: Sets of nonadjacent pacemakers alternate activity. When one set of pacemakers was on, they inhibited all the pacemakers that could inhibit them, and their intrinsic currents provided self-excitation for the duration of their burst. In fact, when the network was in this state, we found that small perturbations such as injecting small depolarizing or hyperpolarizing currents into the pacemakers did not cause it to change state. Once the depolarizing intrinsic current shut off, however, the

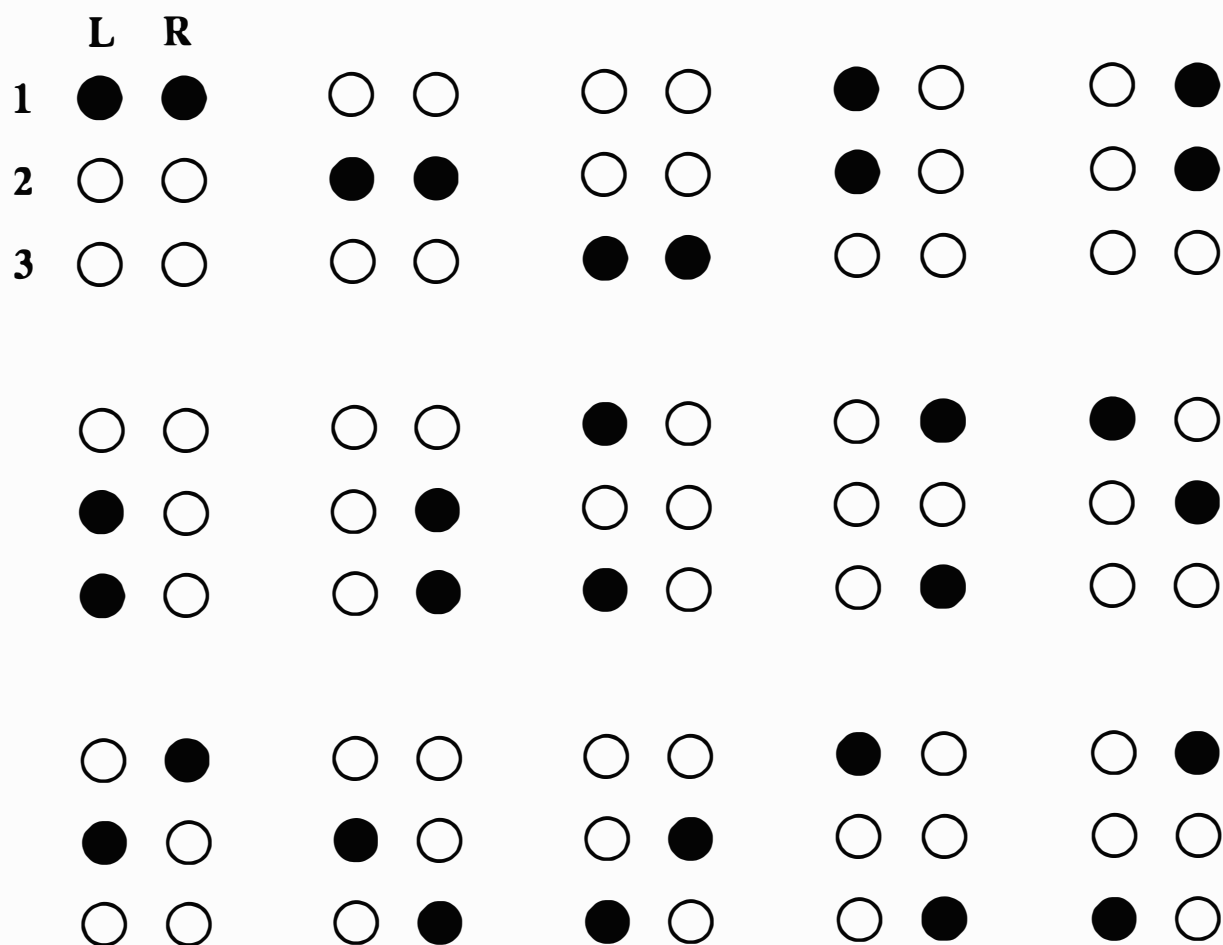


Neural Network for Control of Hexapod Locomotion

Figure 6.1 C, command neuron; P, pacemaker neuron. Subscripts refer to leg (R, right; L, left; 1, front; 2, middle; 3, back).

other triad of pacemakers was released from inhibition and depolarized in response to the steady input from the command neuron. Since their inhibitory intrinsic current was ready to shut off, their depolarizing intrinsic current turned on, and the network moved to its other stable state (figure 6.3; see figure 6.2 for the key to the phase plots). Interpreting the outputs of the pacemakers as motor commands, one would say that the network was generating a tripod gait, in which alternating tripods of legs provide support and push backward, while the other tripod swings forward.

Insects do not use just the tripod gait when they walk. In fact, they show a continuous range of gaits slower than the tripod gait. Graham (1977) developed a model for gait generation based on the idea that coupled relaxation oscillators could be stably entrained, and generate multiple gait patterns. To instantiate this idea in our central network, we altered the connections from the command neuron to the rear pacemakers, so that the amount of current that it injected into the two rear pacemakers was 5% less than the amount it injected into all the other pacemakers. This gradient of excitation changed the intrinsic oscillation rate of the pacemakers, causing the rear pacemakers to oscillate most slowly. As a consequence, their connections to the middle pacemakers were capable of inhibiting these neurons shortly before they would have



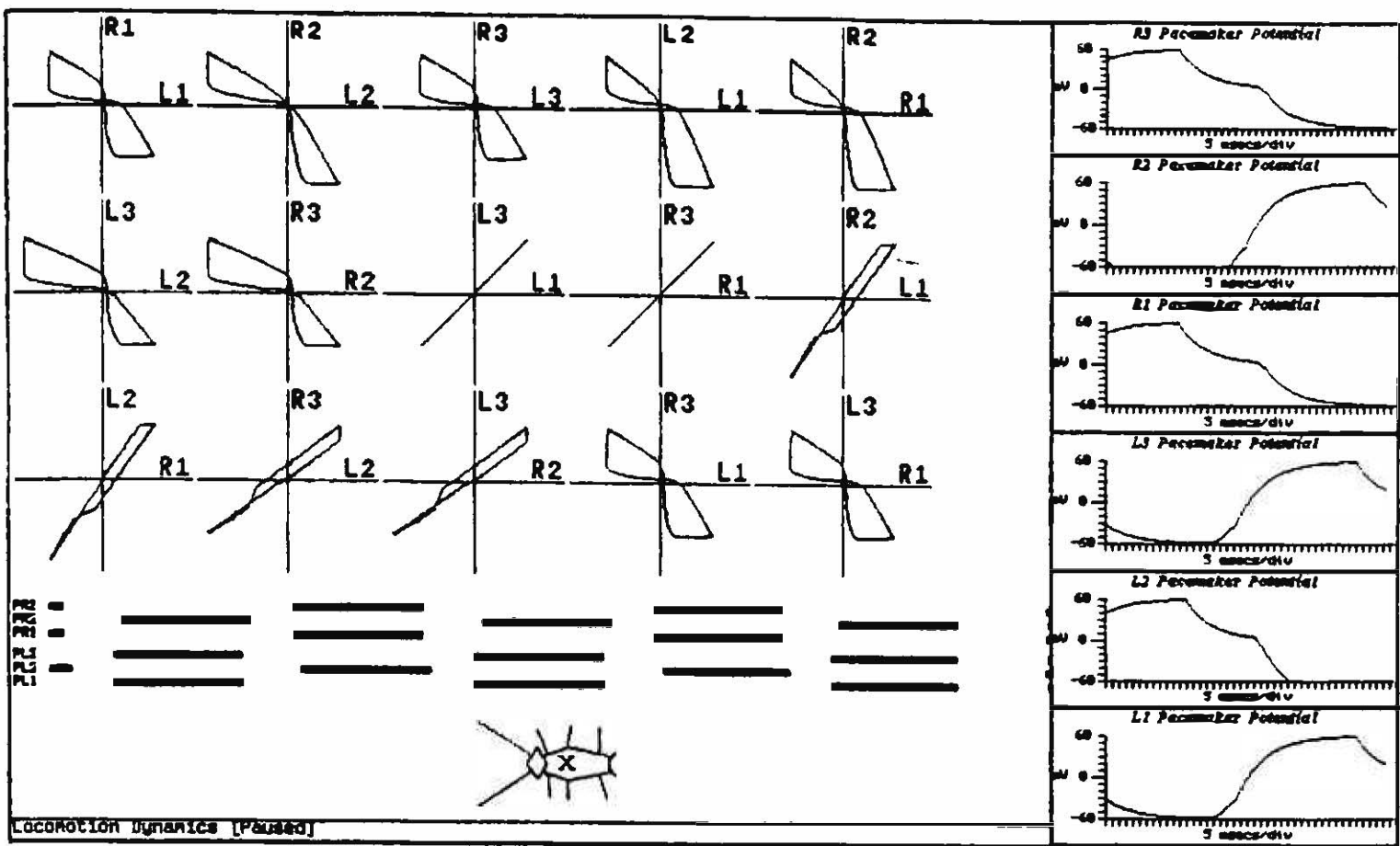
Pacemaker Pairing Patterns Producing Phase Plots

Figure 6.2 Pacemakers whose voltages are plotted against one another in each phase plot of figures 6.3, 6.4, 6.7, and 6.8 are indicated schematically by pairs of darkened circles. Plots numbered consecutively from 1 (upper left) to 15 (lower right).

burst, slowing their burst frequency and entraining them. In turn, the middle pacemakers entrained the front pacemakers. Once this pattern of entrainment was established, pacemakers depolarized in an orderly fashion, from back to front. Changing the excitation level of the command neuron generated a continuous range of similar couplings between the pacemakers (Chiel and Beer 1989).

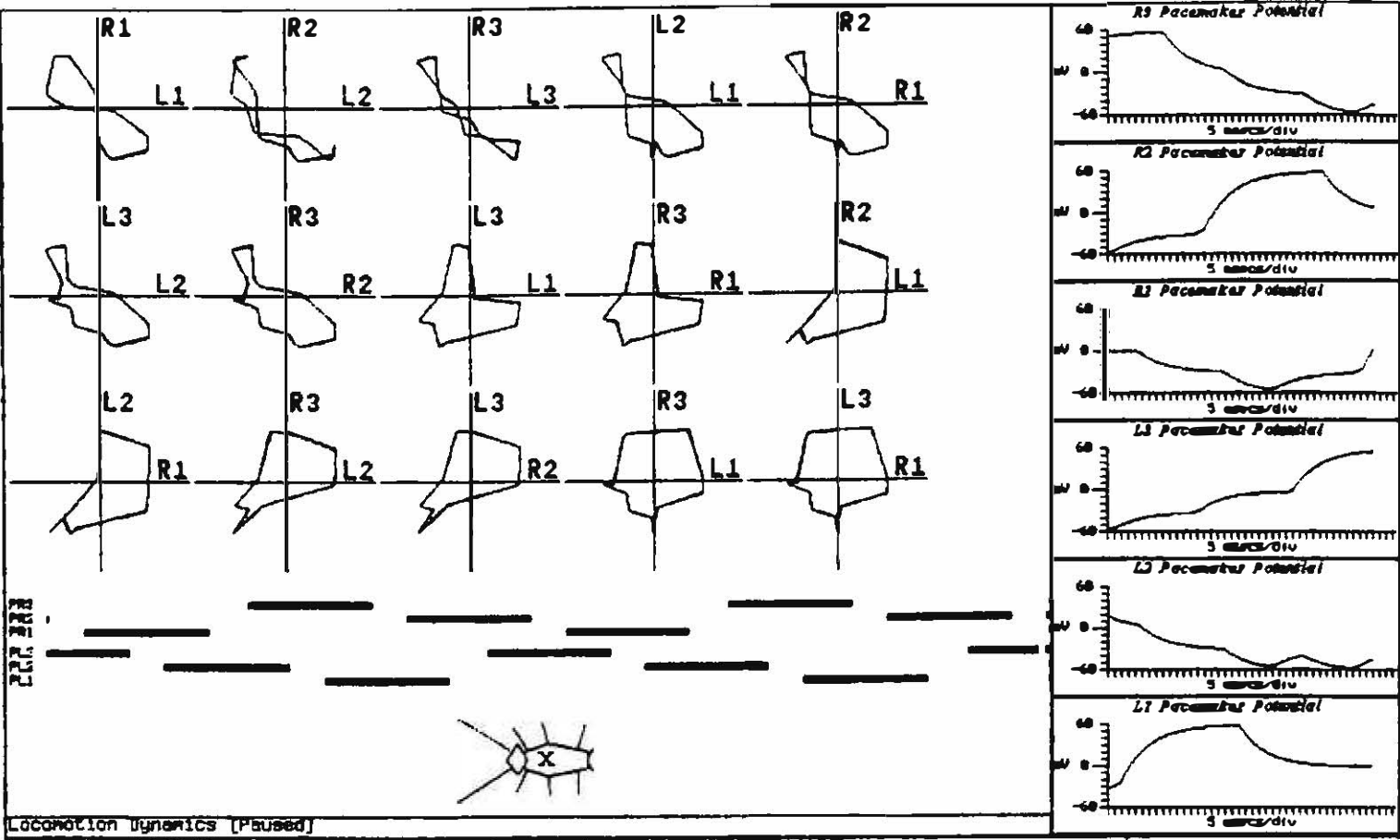
In these slower gaits, fewer pacemakers are on at any time. Thus, the stability of each of the states is not primarily enforced by the connections between the pacemakers, but by their intrinsic currents. The pacemaker that is on is held on by its depolarizing intrinsic current, and those that are off are held off by their hyperpolarizing intrinsic currents, as well as by inhibition from other pacemakers (figure 6.4). In fact, when we shortened the maximum duration of the hyperpolarizing intrinsic current, the network was not capable of generating the slowest gaits.

These results demonstrate that incorporating intrinsic dynamics into individual model neurons, and connecting them in a particular architecture, can generate a neural network that readily produces a continuous range of neuronal outputs that could be utilized for the control of locomotion.



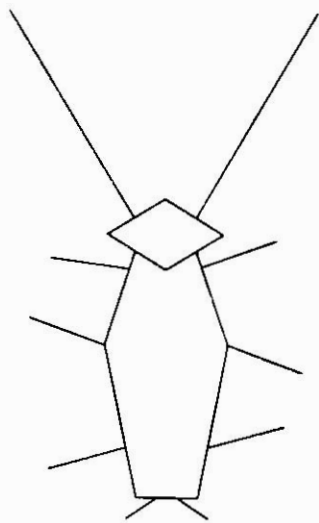
Phase Plot of the Tripod Gait in the Isolated Central Network

Figure 6.3 Here and in figures 6.4, 6.7, and 6.8, pacemaker potentials plotted on the right; 15 phase plots (pacemaker potentials plotted against one another) shown in the upper left of the panel. The lines at the bottom of the panel indicate the times at which the pacemakers were above threshold, thus producing outputs. Data shown here were obtained from the network shown in figure 6.1; command neuron input = 0.5 nA, and simulation time step = 1 ms. Outputs of the command neuron to all pacemaker neurons were identical. Ideally, only two kinds of plots would result: lines at $+45^\circ$ (for nonadjacent pacemakers whose potential varies entirely in phase, e.g., L1, L3, and R2), and lines at -45° (for adjacent pacemakers whose potentials are entirely out of phase, e.g., L1 and L2, L1 and R1). Plots of L3 against L1 (plot 8) and R3 against R1 (plot 9) match this ideal, and all six in-phase plots (plots 8 to 13) are close to $+45^\circ$, while all nine out-of-phase plots (plots 1–7, 14–15) are close to -45° . In all plots, departures from ideality result from two main reasons. One is the greater inhibition received by pacemakers L2 and R2 (three rather than two inhibitory inputs, as shown in figure 6.1). The other is that as one pacemaker begins to turn off and hyperpolarize, the out-of-phase pacemaker remains strongly hyperpolarized, and only turns on after a delay, in part due to its resistance and capacitance. These delays cause the plots to take on a “figure eight” shape.



Phase Plot of a Slower Gait in the Isolated Central Network

Figure 6.4 Connection strengths from the command neuron to the pacemakers for legs R3 and L3 were lowered by 5% compared to connection strengths to the other four pacemakers. Command neuron input = 0.34 nA; simulation time step = 1 ms. Compare to figure 6.3. Pacemakers in the same segment (e.g., L1 and R1) are 180° out of phase with one another (i.e., when one is on, the other is off), in part because of their reciprocal inhibitory connections, and thus their phase plots approximate a -45° line (plots 1–3). Adjacent pacemakers in two different segments (e.g., L1 and L2) are also out of phase with one another for the same reason (plots 4–7), though the phase is 90°. Finally, pacemakers in the first and last segments on either side are out of phase with one another, due to intrinsic inhibition of the third segment pacemaker during first segment pacemaker activity and to inhibition from the contralateral pacemaker (plots 8, 9). A major difference between these plots and those in figure 6.3 is that there are periods of time when both pacemakers are receiving inhibition (both intrinsic and extrinsic); that is, the maintenance of the 180° phasing is not solely due to inhibition from a single source, as it is during the tripod gait; thus, points in these phase plots enter the third quadrant. The remaining phase plots (10–15) show those pacemakers that do not directly inhibit one another, and whose activity partially overlaps. Since these pacemakers are nonadjacent, each tends to change state while the other remains in its current state. The ideal plot of such a relationship would be a box, and these plots in fact enter all four quadrants. Departures from ideality are in part due to the three sources of inhibition received by the middle pacemakers, as opposed to the two sources of inhibition received by all other pacemakers, which now occur at different times.



Kinematic Body Model

Figure 6.5 Kinematic model of a hexapod. See text for details.

Peripheral Plant for Hexapod Locomotion

Our kinematic model of a hexapod was loosely based on the body of the American cockroach, *Periplaneta americana*. It consists of a head, abdomen, two antennae, and six legs (figure 6.5). Each leg has two degrees of freedom: it can swing forward or backward, and it can be raised or lowered. Legs can only generate forces on the body when they are down. Forces exerted by legs in the forward direction are summed, and determine the net force applied to the body. In the simulated environment, it was assumed that the hexapod's motion was highly damped, so that the velocity of the body (rather than its acceleration) was proportional to the force applied to it. Finally, the body as a whole was considered to be statically stable if the polygon formed by the legs that were down contained its center of mass, and unstable otherwise. If instability persisted for more than a simulation time step, the body fell down. As long as the body was fallen, legs could not apply forces to move it forward.

Even this simple kinematic hexapod demonstrated the role that the periphery can play in affecting its neural controller. First, the forces generated by the different legs are coupled through the body, so that the movements of all legs that are down are affected by the other legs that are down. Second, when the hexapod falls down, the relative phasing of the legs can change, since legs that are swinging can continue to move forward, while those which are down cannot. Third, for each leg, there are two sensors that indicate whether the leg has moved far forward or backward, and there are three effectors: one that can swing the leg back, one that can swing it forward, and one that can raise or lower a leg. Thus, the activation of 18 effectors must be properly coordinated with the inputs from 12 sensors to generate locomotion. As a consequence of these properties of the periphery, the properties of the neural network controlling locomotion change when the network is embedded in the model hexapod (see below).

Interactions between Neural Network and Periphery in Gait Generation

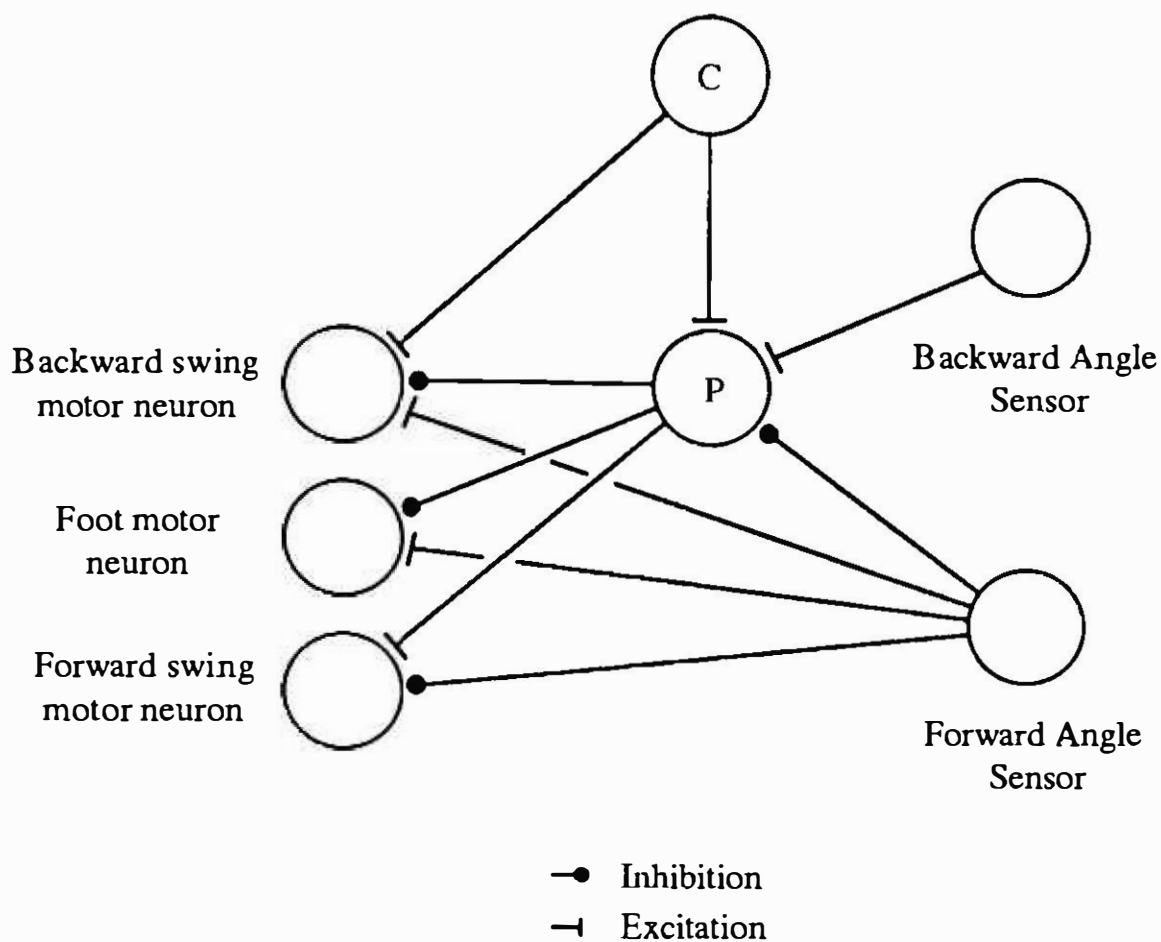
Pearson and his colleagues also obtained data which suggested that the pattern generator for locomotion in the cockroach was strongly shaped by sensory inputs. For example, they found that receptors of the hairplate of the trochanter acted to limit the forward movement of the leg during swing and to facilitate the transition from swing to stance. Similarly, receptors from the campaniform sensilla affected the duration of the stance phase and may serve to compensate for variations in load (Pearson et al. 1973).

These data suggested to us how to embed the central locomotion controller in the kinematic body model (figure 6.6). In contrast to the centralized neural network, the command neuron provided outputs of equal strength to all six pacemaker cells. The pacemakers were linked to the body by providing them with connections to motor and sensory neurons. Three motor neurons for each leg were responsible for activating the effectors which swung a leg forward, backward, or raised or lowered its foot. Two sensory neurons received input from the sensors indicating whether the leg had swung nearly all the way forward, or all the way backward. The single command neuron made additional direct connections to the backward-swing motor neurons for all six legs.

A single cycle of leg movement involves two phases: a *swing* phase, during which the leg swings forward, and a *stance* phase, during which the leg is down, moving backward, and applying force to the body. The swing phase was primarily controlled by the activity of the pacemaker. When the pacemaker was active, it raised the foot by inhibiting the foot motor neuron, and swung the leg forward by activating the forward-swing motor neuron. The stance phase was primarily due to the activity of the command neuron, since it directly activated the backward-swing motor neuron, and the foot motor neuron kept the foot down unless it received inhibition.

The transition from swing to stance was aided by a direct connection from a sensory neuron to the motor neurons, which we termed a *stance reflex*, and has been observed in insects: When the forward angle sensor was activated, indicating that a leg had almost swung all the way forward, it directly inhibited the pacemaker cell, inhibited the forward-swing motor neuron, and excited the backward-swing motor neuron. The transition from stance to swing was aided by the backward-angle sensor, which was activated when a leg had swung almost all the way back, and responded by exciting the pacemaker neuron, thereby initiating a burst in that pacemaker and the onset of the swing phase in that leg.

In order to obtain the gradient of pacemaker periods that is essential for entrainment (as described above), the angle ranges of the rear legs were made slightly longer than those of the middle or front legs. As a



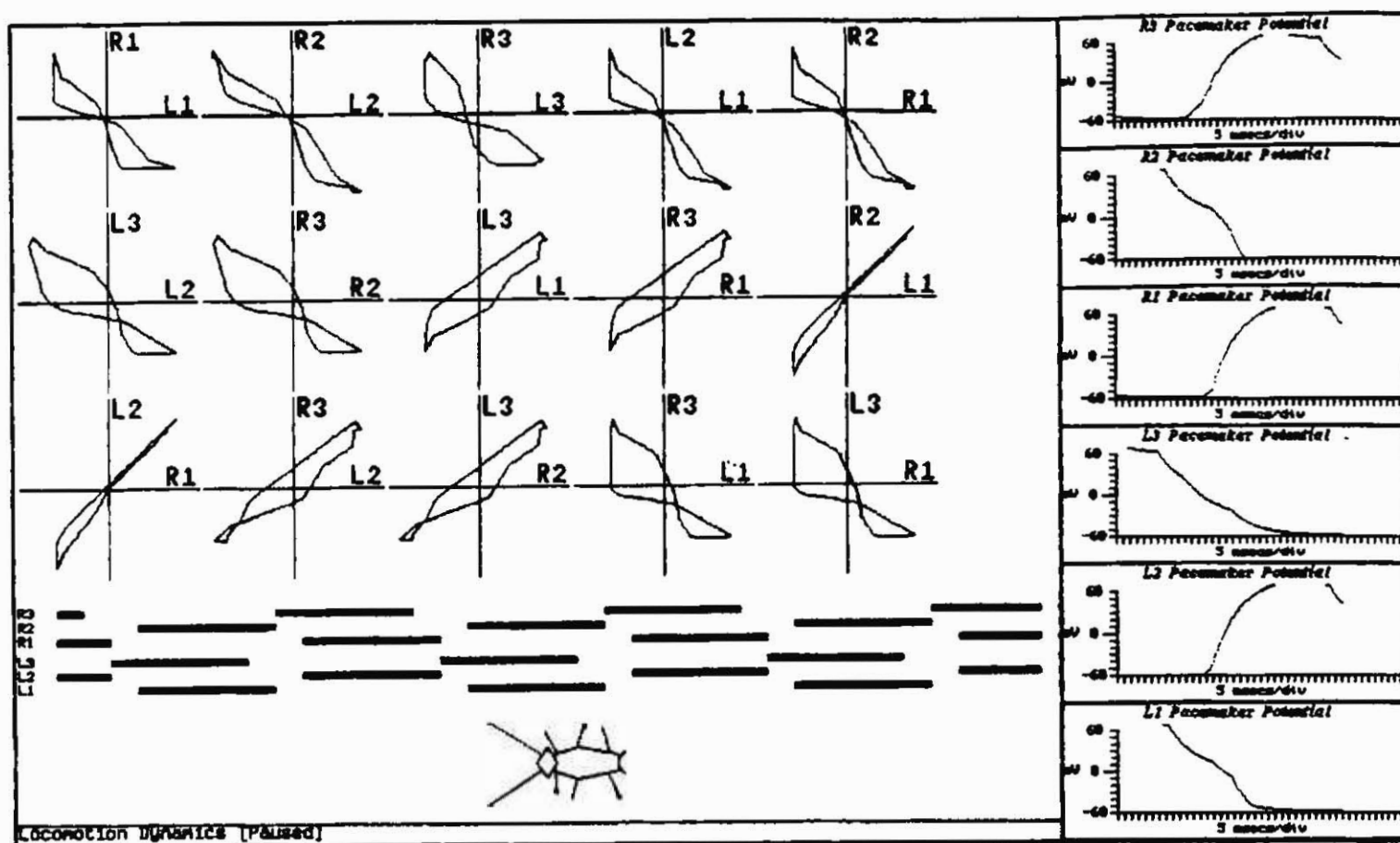
Neural Network Controller for a Single Leg

Figure 6.6 Single controller shown; copies of these sensory and motor neurons were associated with each leg, so that the final controller had a total of 37 neurons (1 command neuron, 6 pacemakers, 12 sensors, and 18 motor neurons).

consequence of this difference in leg angle ranges, the rear pacemakers cycled at a slower rate than either the middle or front pacemakers. In turn, this led to stable patterns of entrainment, so that rear legs swung before middle legs, which swung before front legs, forming a *metachronal wave* of swings that has been observed in insects.

As activation of the command neuron was increased, the simulated insect generated a continuous range of statically stable gaits, resembling those seen in insects (Wilson 1966). Since it can generate a continuous range of such gaits in the absence of sensory inputs, what role did the sensory inputs play in the functioning of the controller? To address this question, we used the technique that neurobiologists have traditionally utilized for the analysis of neural circuitry—we lesioned parts of the controller, and observed the effects of the lesions on its behavior (Chiel and Beer 1989).

Our first observation was that the controller was extremely robust. Single lesions of sensors or of connections between pacemakers did not prevent the simulated insect from walking stably over a wide range of speeds. We found, however, that the role of sensory inputs changed as the speed of the gait varied. At the fastest gait, the tripod gait, we found that all sensory inputs could be eliminated, and the insect would still walk stably. Based on our analysis of the tripod gait described above, we deduced that sensory inputs could help to reinforce the tripod gait,

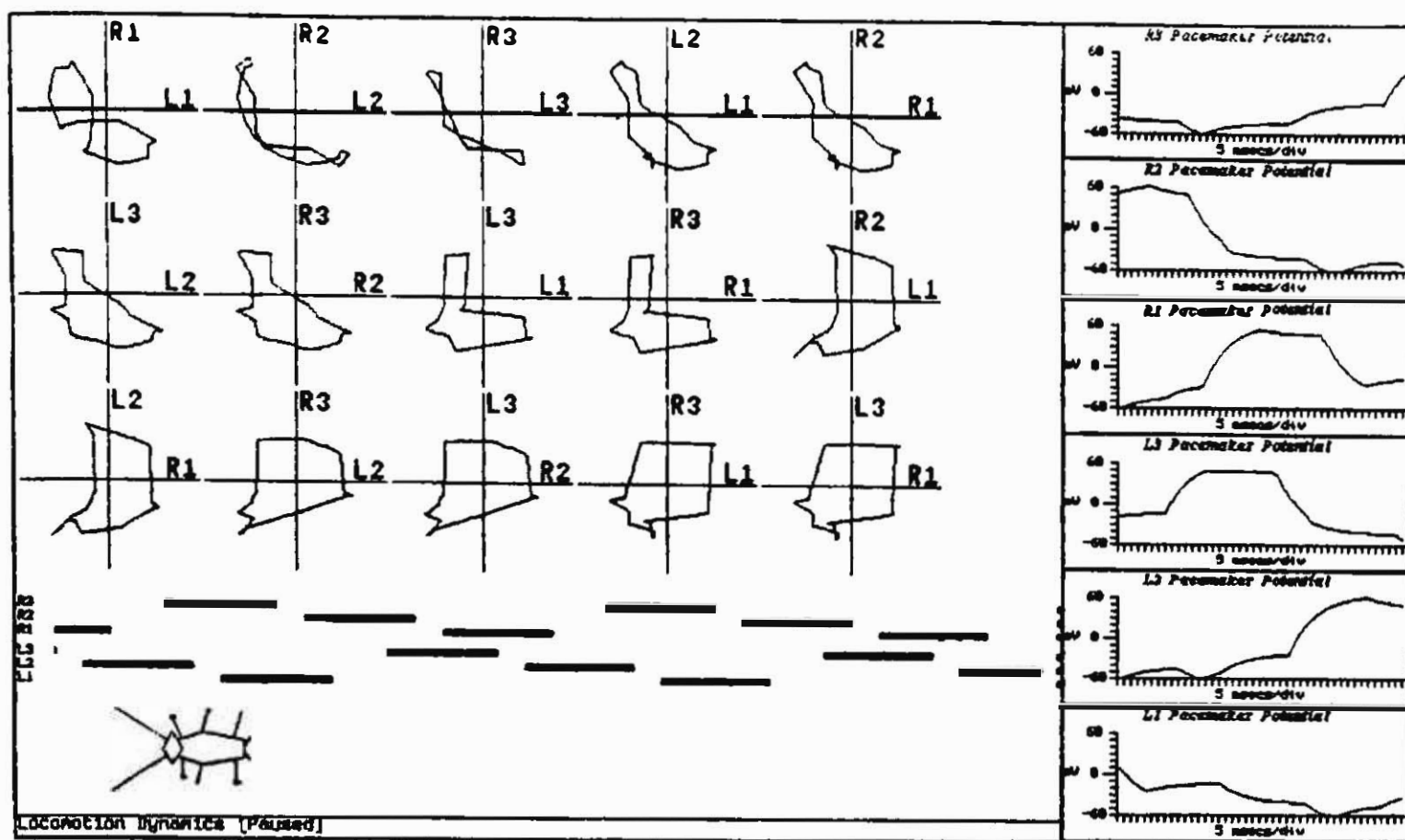


Phase Plot of Tripod Gait in Intact Network

Figure 6.7 Since this network was connected to motor outputs, stepping patterns are shown on the bottom left of this figure and in figure 6.8 (a black bar indicates when a leg is swinging forward). Command neuron activation = 0.5 nA; simulation timestep = 1 ms. Phase plots are remarkably similar to those of the tripod gait in an isolated network (figure 6.3). Nonadjacent pacemakers (in phase, plots 8–13) have plots that are close to $+45^\circ$ lines. Adjacent pacemakers (out of phase, plots 1–7) have plots that are close to -45° lines. Although the last two plots (14 and 15) are of nonadjacent pacemakers, they fall into two different tripods and have plots that are close to -45° lines. The main effect of sensory inputs can be seen by focusing on the maximum and minimum values taken by the pacemakers; they have larger absolute values than in figure 6.3, reflecting the excitatory and inhibitory effects of the backward and forward angle sensors, respectively (figure 6.6). Thus, the sensory inputs act to enforce the pattern generated by the central neural network, but are not essential for generating this pattern of activity.

and if they occurred at incorrect times, they could disrupt it, but that the gait was essentially established and maintained by central connections (figure 6.7).

In contrast, certain sensory lesions had devastating effects on slower gaits. For example, lesioning both backward-angle sensors for both back legs had no effect on the tripod gait. However, the identical lesion was extremely disruptive for the slower gaits. Since the backward-angle sensors are crucial for establishing the rate of pacemaker cycling, disrupting them affects the entrainment that gives rise to the metachronal wave. During the slowest gaits, a pacemaker that is on is kept on primarily by its intrinsic depolarizing current, and the ones that are off are kept off primarily by their intrinsic hyperpolarizing current. An additional constraint is that, during the time it is on, the active pace-



Phase Plot of Slower Gait in Intact Network

Figure 6.8 To produce the slower gait, the command neuron was injected with -0.25 nA, with simulation timestep = 1 ms. Compare to slower gait production in the isolated network, shown in figure 6.4. There is considerable similarity between the two figures, but in figure 6.4, the slower *activations* of the rear pacemakers (due to the reduced command input) lead to the appropriate couplings between the pacemakers to enforce the slower gait. If this change is not made to the isolated network, it cannot generate slower gaits. In contrast, in the intact network, the slower *sensory inputs* to the rear pacemakers lead to these couplings, since command neuron inputs to all pacemakers are identical. In addition, the sensory inputs lead to larger and more rapid activations and inactivations of the pacemakers, which can be seen by comparing the slopes of the lines in the phase plots (e.g., compare plot 15 here to the equivalent plot in figure 6.4).

maker should not receive inhibitory inputs from the forward-angle sensor, while those that are off should not receive excitatory inputs from their backward-angle sensors. These requirements place constraints on the rates at which the sensory inputs could come into the network without disrupting its function. In fact, we found that activating sensors at the wrong time had devastating effects on the function of the controller. Thus, the formation and maintenance of slower gaits is crucially dependent on sensory input, in contrast to the tripod gait (figure 6.8). An interesting parallel to these results is found in insects. Proprioceptive inputs seem to play a much larger role in maintaining slower gaits, but less of a role in the tripod gait (Zill and Moran 1981).

This analysis of the locomotion controller serves as a striking illustration of the perspective we have described. It illustrates that the ongoing interactions between the periphery and a neural network controller are central to the generation of adaptive patterned behavior, which in this case is a continuous range of statically stable hexapod gaits.

Peripheral and Central Interactions in the Control of a Hexapod Robot

In order to determine whether the controller for locomotion developed in simulation could control an actual robot, despite that robot's inertial properties, delays, and other differences from the simple simulated hexapod, we have applied our controller to a hexapod robot developed by two colleagues at Case Western Reserve University, Roger Quinn and Kenneth Espenschied of the Department of Mechanical Engineering (Quinn and Espenschied 1993). Each leg of the hexapod robot had two degrees of freedom—one that allowed it to extend or contract radially, and a second that allowed it to swing backward or forward.

In attempting to use our controller in the hexapod robot, we discovered that delays in positioning the legs could disrupt its function. In order to utilize the velocity commands generated by the controller, we had built a position controller for each leg. It was based on the equilibrium point principle demonstrated by Bizzi and his colleagues (Bizzi et al. 1991; see above): Perturbations of the leg from the equilibrium position generated a restoring force that returned the leg to that position. Initially, when the robot was directed to walk using the locomotion controller described above, we observed that the metachronal wave had inexplicable gaps in it, and that the phasing of leg swings across the body (180° out of phase in the original controller) was also disrupted. What was different in the robot?

Plotting the desired leg position against the actual leg position revealed two differences between the simulation and the robot. First, it showed that there were delays, because the position controller could not instantaneously move the leg to its new position. Second, it showed that the delays were asymmetric. These asymmetries were due to inherent differences in the electronics for each position controller, as well as differences in the motors for each leg. We found that we could cause identical deficits in the simulated controller by introducing delays. Delaying all legs caused gaps in the metachronal wave; asymmetric delays disrupted the phasing of leg swings across the body. When we retuned the components of the robot to compensate for these delays, it largely corrected the problems (Beer et al. 1992). Thus, constructing the robot and tuning the controller to work with its periphery revealed new aspects of the controller to us. In addition, we found that the controller continued to be extremely robust in response to lesions, and that the robot was capable of walking stably even after single sensory, central, or motor connections had been removed (Chiel et al. 1992a).

Summary: Controller for Hexapod Locomotion

The locomotion controller demonstrates the significance of several of the main themes we have developed in this chapter. First, it demonstrates the importance of incorporating biological reality into artificial

neural networks. The richness and appropriateness of the dynamics of this neural network are due to the intrinsic properties of the pacemaker cells and the carefully designed architecture of interconnections between them. Second, it demonstrates our central thesis that patterned motor behavior emerges from the ongoing interactions between a neural network and the periphery that it controls. Thus, the range of gaits produced by the controller emerges from the interactions between sensory input and the central neural network, and this in turn causes the controller to be extremely robust to perturbations that would completely incapacitate a more centralized control architecture. Finally, the locomotion controller demonstrates that it is possible to take ideas from biology, test them in simulation, and then use them for developing specific engineering applications, such as the control of hexapod locomotion in robots. In turn, constructing these devices provides new insights into the nature of the controller (such as its sensitivity to asymmetric delays) that could provide the impetus for new experimental work in the biological system.

CONCLUSIONS

The underlying thesis of this chapter has been that the coupling between the dynamics of a neural network and the periphery that it controls is essential for the flexibility and adaptability of the behavior of biological organisms. This broader perspective suggests conclusions that should be of interest to biologists and to engineers. It also suggests new directions that need to be pursued in both of these areas.

Issues for Neurobiology

Neurobiologists are intimately familiar with the bewildering complexity of biological nervous systems. In addition, many have become increasingly aware of the complexity of the coupling between the nervous system and its periphery. We hope that this chapter has indicated the extent to which many properties of the nervous system, and even individual neurons, may not be comprehensible in the absence of the periphery. What is more, many properties may not be experimentally observable unless the nervous system is interacting with its periphery. In the absence of proprioceptive feedback, modulatory influences from the periphery, and so on, a neural network may adopt a very different configuration.

Our studies of the short-term couplings between the locomotion controller and the periphery during different walking gaits illustrate these points. There has been an ongoing debate over the last few years as to whether there is such a thing as a central pattern generator, a neural circuit that in the absence of any sensory information can generate a complete pattern of behavior. Pearson has argued that no such thing exists; all

pattern generators are shaped by sensory feedback (Pearson 1985); other features of central pattern generators are discussed in chapters 5 and 7. The lesion studies of the locomotion controller can be utilized to address this point, although it cannot be claimed that the locomotion controller is a faithful model of biological reality. The locomotion controller utilizes both central and sensory information in order to generate the different gaits. In different contexts, these sources of information are more or less essential. Thus, during high-speed gaits, central information seems to be most significant; in contrast, during low-speed gaits, sensory information seems to be most significant. In all gaits, however, both sources of information contribute. Although removing them may not prevent the system from functioning at all, it will degrade its performance.

The work reviewed in this chapter suggests some potentially productive areas that should be explored by neurobiologists. A focus on the coupling between the nervous system and its periphery is likely to provide new insights into the function of both. Given the limitations of current tools, it may be most prudent to focus on simpler systems, in which the peripheral plant and the neural network controlling it can be most thoroughly analyzed experimentally.

A conceptual framework that we would propose for this area of investigation has been termed *computational neuroethology*, the simulation of the adaptive behavior of simpler whole animals as they interact with their environment (Beer 1990). This approach requires that an investigator synthesize knowledge of neural mechanisms, an organism's periphery, and its environmental niche in order to study the mechanisms of adaptive behavior. A dynamic systems perspective may be extremely fruitful for the understanding of autonomous agents more generally (Beer 1992). We believe that these approaches are likely to prove extremely fruitful in the coming years.

Issues for Engineering

In the last few years, engineers have discovered that artificial neural networks may provide a new approach for solving a variety of different problems, such as pattern recognition. It is still unclear how useful such a technology will be for problems that require high degrees of accuracy and precision. It may also be true that any particular biological design could be improved by clever engineering.

Nevertheless, in this chapter we have discussed several ideas that may have engineering utility. As we have attempted to illustrate in the body of the chapter, the complexity of biological neural networks may be used to advantage for engineering. Thus, the heterogeneous neural network for locomotion described in this chapter has been utilized to control a six-legged robot. Unlike other controllers for hexapod locomotion, the controller has no centralized gait generator; rather, the different gaits

are emergent properties of the coupling between the different pacemakers, their intrinsic properties, and the sensory inputs to them. As a consequence of this design, the controller is extremely robust to lesions that would cripple more centralized controllers.

The peripheries of biological organisms also suggest new ideas for actuator devices. Muscular hydrostats and hydrostatic skeletons may provide design ideas for actuators that can manipulate highly irregular objects, and penetrate into extremely irregular crevices (Wilson 1984). With the advent of contractile biopolymers, it may become feasible to construct such devices. Understanding the neural control paradigms used by biological organisms for these structures will be essential for exploiting their full power.

The mechanisms of coupling between nervous systems and their periphery also suggest ideas for engineering. For example, the equilibrium hypothesis and the experimental work of Bizzi and co-workers suggest new ideas for the controls of multijointed actuators with excess degrees of freedom. Indeed, simulations have shown that this approach can work successfully (Bizzi et al. 1991). More generally, these results suggest that, in tackling the control of autonomous systems, considerable thought should go into the design of the peripheral plant. By changing the properties of the periphery, it may be possible to enormously simplify aspects of the control problem.

Finally, the traditional engineering sequence for devising new computing devices has been design, construction, and then programming or training. Indeed, the appeal of artificial neural networks and the learning algorithms that have been devised to train them has been that the new technology can be applied to control and pattern recognition problems in ways that are quite similar to more traditional engineering approaches. However, the drawback of such an approach is that much of the complexity of biological neural networks has been sacrificed, along with the richness and power that they bring to the solution of control problems. How would it be possible to retain more of this richness and complexity?

In the biological realm, new ecological niches open, and organisms evolve to fill them. They are not constructed, but develop, and the architecture and plastic properties of their nervous systems allow them to function in the new ecological niche. Thus, the biological processes of *evolution*, of *development*, and of *plasticity* in the midst of defined neural architectures may serve as the basis for new ideas for the design of computing devices. A recent example is provided by the work of Beer and Gallagher (1992), who have shown that it is possible to use genetic algorithms to evolve recurrent neural networks that can control chemotaxis or hexapod locomotion. By developing artificial analogs of development and evolution for the design of neural networks, it may not be necessary to place as heavy a burden on algorithms for training artificial neural networks, and it may be possible to incorporate much more of

the richness and complexity of biology in a readily useable fashion. If a concerted effort is made to develop such approaches, biological inspiration may have a very significant impact on engineering in the years to come.

ACKNOWLEDGMENTS

H.J.C. thanks the NSF and the NIH for supporting some of the work described in this chapter with grants BNS-8810757 and HL-25830-11A1. R.D.B. thanks the ONR for supporting some of this work with grant N00014-90-J-1545. Both authors are grateful for support from the Cleveland Advanced Manufacturing Program (CAMP), and the Center for Automation and Intelligent Systems Research at CWRU.

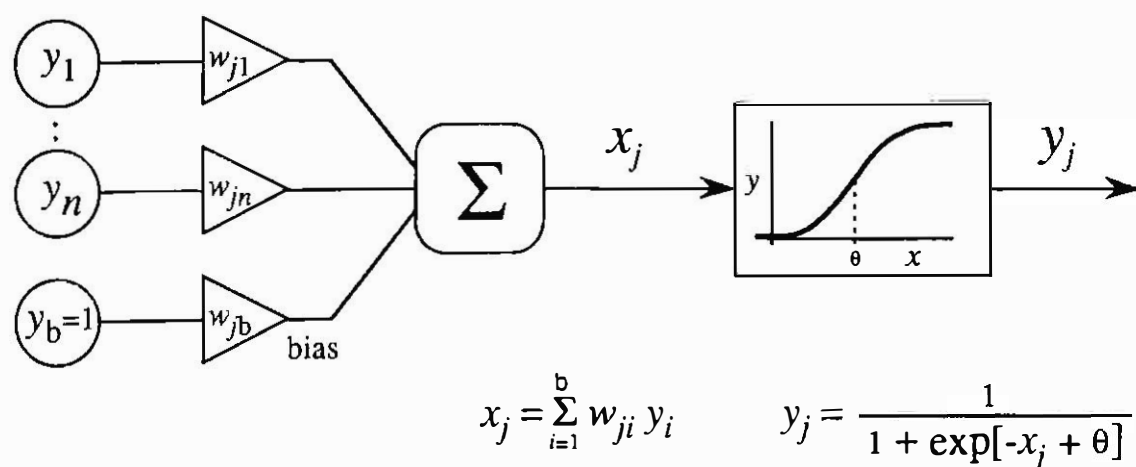
Dynamic Neural Network Models of Sensorimotor Behavior

Eberhard E. Fetz

The goal of systems neurophysiology is to understand the neural mechanisms underlying behavior. Traditional approaches toward this goal have involved anatomical and physiological techniques that reveal the connectivity or activity of specific components of the system. Yet, despite innumerable studies documenting how lesions affect behavior, how stimulation evokes movement, and how single units fire in behaving animals, we still have no consensus about how biological neural networks actually generate movement. The basic reason that the experimental data have not yet converged on a definitive causal explanation of motor behavior is simple: The traditional techniques inevitably provide only limited samples of the system. For example, recordings of single unit activity indicate their response patterns, but nothing about their relation to the rest of the system. What is missing is what artificial neural network models can provide: a method of generating working models of the complete system.

Two general approaches to neural network modeling should be distinguished at the outset. The bottom-up approach involves synthesizing experimental details about cellular mechanisms into a working network model. This approach appeals to physiologists concerned with preserving biological reality, who wish to demonstrate how physiological neurons could function when combined into circuits; however, the bottom-up approach runs into problems dealing with the whole organism, since many parameters of the nervous system are largely unknown and must be arbitrarily assigned. Moreover, modeling the whole system with “realistic” neurons soon becomes prohibitively complex. Alternatively, recent advances in neural network algorithms now make it possible to derive top-down simulations of sensorimotor behavior in networks of appropriately chosen connectivity. The latter techniques provide neural networks that perform specified behaviors, derived entirely on the basis of examples of the behavior itself.

The first examples of such top-down simulations involved feedforward networks trained with backpropagation to transform spatial patterns. For example, feedforward networks have been trained to classify the shape of a surface from shaded images (Lehky and Sejnowski 1990).



Input-output Properties for Sigmoidal Units used in Dynamic Recurrent Networks

Figure 7.1 Each unit j generates a continuous activation y_j as output. Its net input x_j is the weighted sum of activities of its input units y_i and the bias, weighted by synaptic connection strengths w_{ji} . This sum is transformed by the sigmoidal input-output or “squashing” function f to limit the output: $0 < y_j < 1$. The offset θ reduces output to near zero in the absence of input.

Another such network combined the retinal coordinates of a visual stimulus with eye displacement to provide a code for the position of the stimulus in head coordinates (Zipser and Andersen 1988a). Interestingly, the hidden units in such networks often have response properties resembling those seen in biological neurons recorded in animals. The shape-from-shading networks developed hidden units with center-surround receptive fields like visual cortex cells, even though the simulated task did not explicitly require such properties or present stimuli with edges. The reasons that units in network simulations of a task often have such close correspondence with units in biological systems may be more than coincidental; it may be related to basic properties of identification models, as discussed by Zipser (1992).

The supervised training procedures have now been extended beyond feedforward networks, which transform spatial patterns through serial layers, to dynamic recurrent networks, which can deal with spatiotemporal patterns and have feedback connections. Results obtained from top-down simulations of behavior with dynamic recurrent networks are the subject of this chapter.

Three key properties distinguish dynamic recurrent networks from other modeling approaches:

First, the units are *dynamic*, meaning they can exhibit time-varying activity. The activation of each unit is a temporal variable, which can represent the firing rate of neurons, their synaptic potentials or some relevant time-varying sensory or motor parameters. This means that the models can incorporate physiologically recorded activity as activation patterns that the network can be required to generate.

Second, the networks have a *recurrent* architecture, which allows unrestricted connectivity. In addition to feedforward connections between successive layers of units, the networks can have feedback and

cross-connections. Thus, the models can incorporate recurrent anatomical pathways.

Third, the networks that simulate a particular sensorimotor behavior can be derived from examples of the behavior by *gradient descent* methods such as backpropagated error correction. The resulting network models provide complete neural network solutions of the behavior, insofar as they determine all the connections and activations of the units that simulate the behavior.

Neural networks that emulate particular sensorimotor behaviors—i.e., that transform spatiotemporal inputs to appropriate outputs—can be derived by various training strategies. These networks usually comprise the units illustrated in figure 7.1, often named “sigmoidal” or “logistic” for the fact that the unit’s output y is a sigmoidal function of its input x . The unit’s input consists of the summed output activation of all other cells connected to that unit times their synaptic weight; a steady input may also be provided by a bias element:

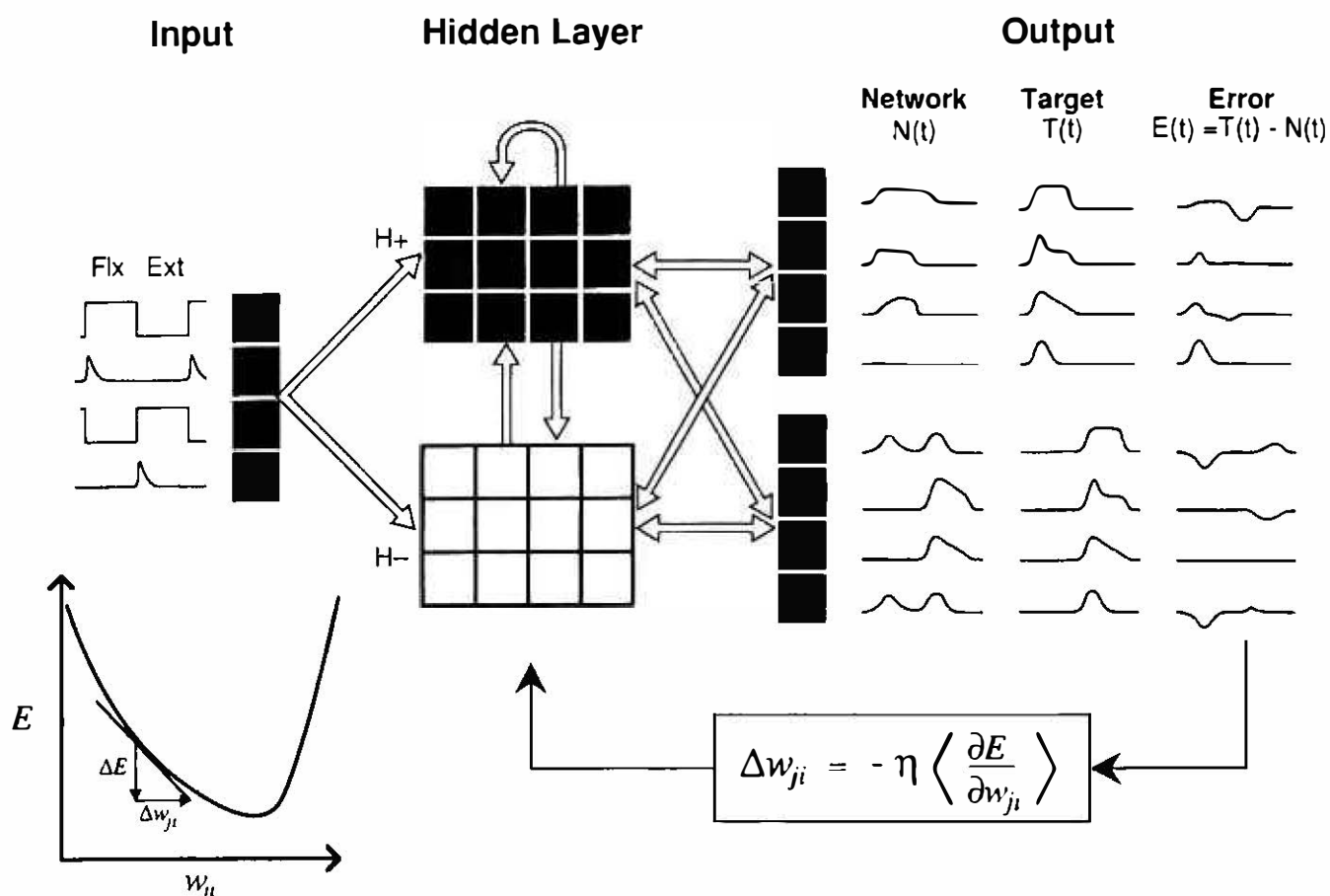
$$x_j = \sum_{i=1}^b y_i w_{ji} \quad (7.1)$$

(see also equation 3.1). The sigmoidal input-output function f mimics a biological neuron’s property of saturating at maximal rates for high levels of input, and decreasing to zero at low input levels:

$$y_j = f(x_j) = \frac{1}{1 + e^{[-x_j + \theta]}} \quad (7.2)$$

Additional properties of these units are provided in chapter 3; chapter 5 describes not only the units but also the input-output function. The function f often includes an offset term θ to ensure that units generate negligible output in the absence of input.

To illustrate the training procedure, figure 7.2 shows a representative network of such units, with network input and output patterns that simulate the step-tracking task described below. The four input units carry signals representing the changing target locations; these signals are transformed to eight output activation patterns, representing the firing patterns of typical motor units in the muscles of a monkey tracking the target. To train the network, the synaptic weights are initially assigned randomly, and the output response of the network is determined; the difference between network output patterns $N(t)$ and the desired target output activations $T(t)$ is the error $E(t)$. The backpropagation algorithm modifies the weights in such a way as to optimally reduce this error. As shown by the equation, the optimal weight change can be computed from the error when the functions are differentiable. As illustrated by the inset in figure 7.2, this weight change implements a *gradient descent* of the error as a function of the weight. After each weight correction the input patterns are presented again and the new outputs



Typical Network Architecture and Training Procedure Used with Dynamic Recurrent Networks

Figure 7.2 This network was used to simulate the step-tracking task (see figure 7.5 for a specific example). The network input consists of four representations of the step target position and target change; the output represents the firing patterns of eight representative motor units in flexor and extensor muscles. The intervening hidden units consist of excitatory and inhibitory groups, with distributed connections indicated by the arrows. Network training proceeds by calculating the difference between the network output $N(t)$ and the desired target activations $T(t)$, and changing the connection weights in such a way as to reduce the error $E(t)$. Inset at lower left illustrates the error as a function of one weight, and the use of the gradient of this function to determine the appropriate weight change.

determined. This process of changing the weights in proportion to the error gradient is repeated iteratively until the network converges on a solution with minimal error.

Several types of algorithms have been used to implement such gradient descent methods for dynamic recurrent networks. Watrous and co-workers derived backpropagation techniques to train dynamic networks to identify acoustic features from speech data (Watrous and Shastri 1986). Their temporal flow model involved a hidden unit layer between input and output that had recurrent connections. Pearlmutter (1989) developed an algorithm for recurrent networks of units carrying continuous temporal signals, and applied this to training a network to generate trajectories in phase space. Williams and Zipser (1989a, 1990) provided a real-time recurrent learning (RTRL) algorithm for training recurrent networks carrying time-varying signals. RTRL updates the weights after each time step, rather than after the end of a sequence, allowing the training to occur in real time, with less storage require-

ments. A useful variant of this algorithm is teacher-forced learning, in which the target activations are substituted for the actual unit outputs during training; this has proven more effective for training networks to generate oscillations. A comparison of various training methods for recurrent networks is presented in Williams and Zipser (1990).

APPLICATIONS

The applications for these dynamic recurrent networks can be classified into three broad categories:

Pattern recognition applications involve identification of spatiotemporal input patterns into discrete categories. A set of input units receiving time-varying signals can be considered to represent a spatiotemporal pattern. Pattern recognition applications include speech recognition (Watrous and Shastri 1986), conditional delay tasks (Williams and Zipser 1989a, b) and recognition of finite state grammars (Smith and Zipser 1989).

Pattern generation networks produce temporal patterns in one or more output units, either autonomously or under the control of a gating input. These include oscillating networks (Pearlmutter 1989; Williams and Zipser 1989a, b) and simulations of central pattern generators (Rowat and Selverston 1991; Tsung et al. 1990), described below.

Pattern transformation networks convert spatiotemporal input patterns into spatiotemporal outputs. Examples include simulations of the leech withdrawal reflex (Lockery and Sejnowski 1990, 1992), step-target tracking in the primate (Fetz and Shupe 1990; Fetz et al. 1990), the vestibulo-ocular reflex (Anastasio 1991a, b; Arnold and Robinson 1991; Lisberger and Sejnowski 1992a), and short-term memory tasks (Zipser 1990, 1991). Recurrent networks have also been trained to simulate analytical transforms such as integration and differentiation of input signals (Munro et al. 1991) and to emulate special-purpose Turing machines (Williams and Zipser 1989b).

Given the generality and power of dynamic recurrent networks, it seems remarkable that the applications are still sufficiently few to summarize in the scope of this chapter.

Oscillating Networks

Biological systems provide numerous examples of autonomously generated periodic motor activity: locomotion, mastication, respiration, etc. The neural circuitry underlying cyclic periodic movements has been dubbed a central pattern generator (CPG). Additional reviews of CPGs are discussed in chapters 5 and 6. Besides motor activity, oscillatory activity of various frequencies also appears in cortical circuitry during the alpha and gamma waves of neocortex and the theta waves of hippocampus. Networks that generate oscillatory activity have been con-

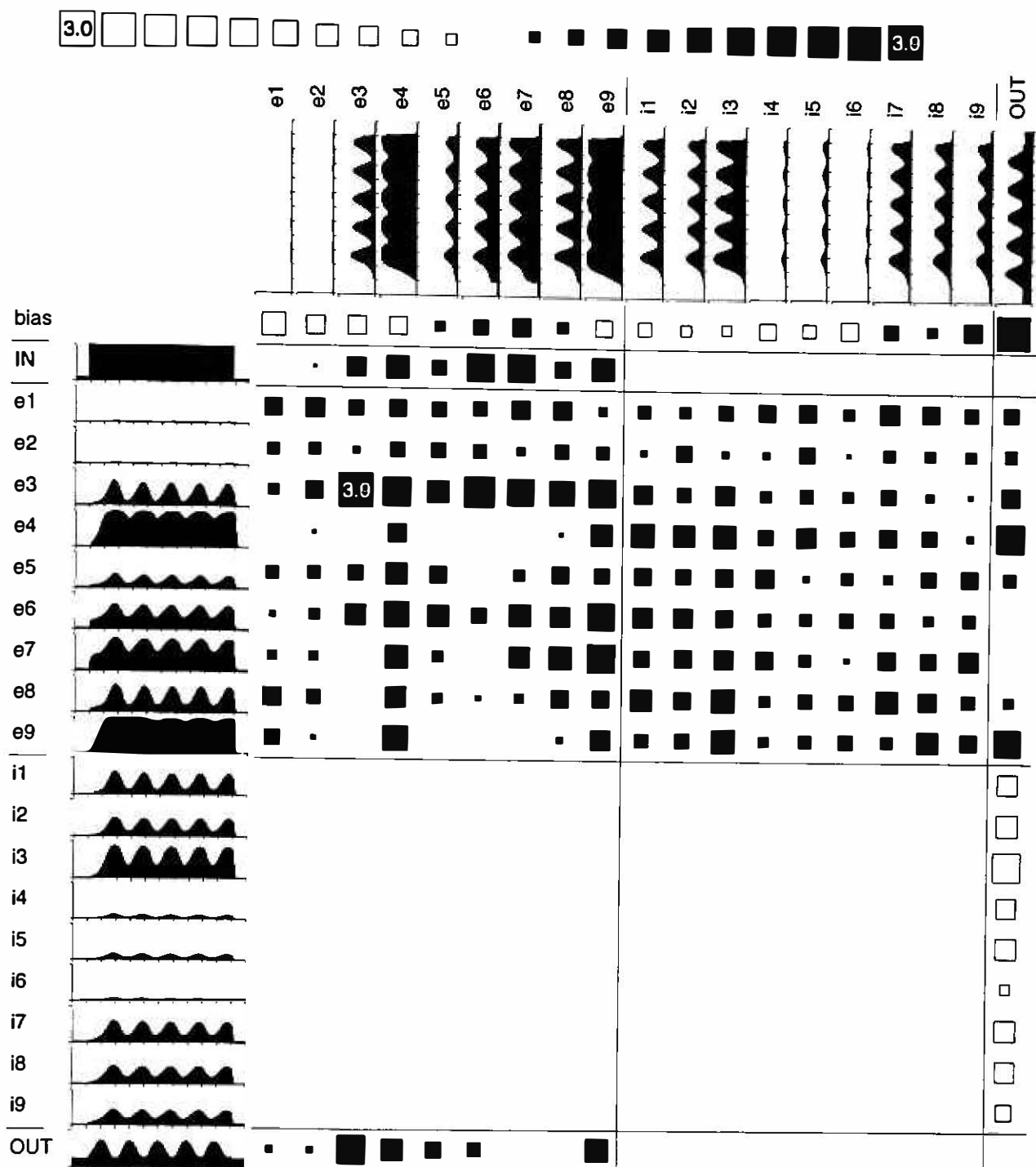
structured with bottom-up strategies (e.g., Buchanan 1992; Grillner et al. 1991), but our interest here is in networks derived by backpropagation techniques.

Williams and Zipser (1989a, b) trained dynamic networks of sigmoidal units to generate oscillatory activity with various specified frequencies. Boolean oscillations between two binary states were readily achieved, but only with teacher-forced training. Dynamic networks could also generate sine-like oscillations with periods of 25 timesteps or less. The smallest circuit that sustained quasi-sinusoidal oscillations consisted of two interconnected sigmoidal units. The frequency of the free-running network was about 10% lower than the trained frequency sustained by the teacher signal. Selverston and co-workers (Tsung et al. 1990) trained pairs of sigmoidal units to oscillate with specific phase shifts. The two units preferred to oscillate with similar (but not identical) phase; training them to exhibit larger phase shifts involved increasing numbers of training cycles. A 15-unit interconnected network without sign constraints could readily be trained to oscillate with equidistant phases (i.e., phase shifts of $2\pi/15$, or 24°).

To simulate the oscillatory activity of neurons in the lobster gastric mill circuit, Tsung et al. (1990) trained a network with the connectivity and sign constraints of known neurons in the mill circuit; this network replicated the correct phase relations of the oscillating interneurons. After its activity was perturbed, the network reverted to the original pattern, indicating that the weights found by the learning algorithm represented a strong limit cycle attractor.

Rowat and Selverston (1991) have developed learning algorithms to train oscillations in dynamic recurrent networks of units with biological properties like gap junctions and membrane currents. They used a generalization of the Williams and Zipser teacher-forced learning procedure to train networks to simulate known oscillatory activations and to observe sign and magnitude constraints on connection weights.

My colleague Larry Shupe and I have also derived autonomous oscillators through dynamic recurrent networks. The input was a gating step that turns network activity on, and the target output was activity oscillating at a prescribed frequency. Figure 7.3 shows a typical recurrent network that generates gated oscillations. The discharge pattern of each unit is shown along the left, next to the row of output weights of that unit. From top to bottom these include the bias (a constant), the gate signal (IN), nine excitatory hidden units (e1–e9), nine inhibitory hidden units (i1–i9) and the oscillating output (OUT). The hidden units and output are shown again at the top, above the column of weights representing the input connections to each unit. The squares in the matrix symbolize the strength of the synaptic connection from the row unit to the column unit. Black squares designate excitatory weights and open squares represent inhibitory weights. (Note that this convention is opposite from that of the Hinton diagrams in chapter 5.) The size of each square

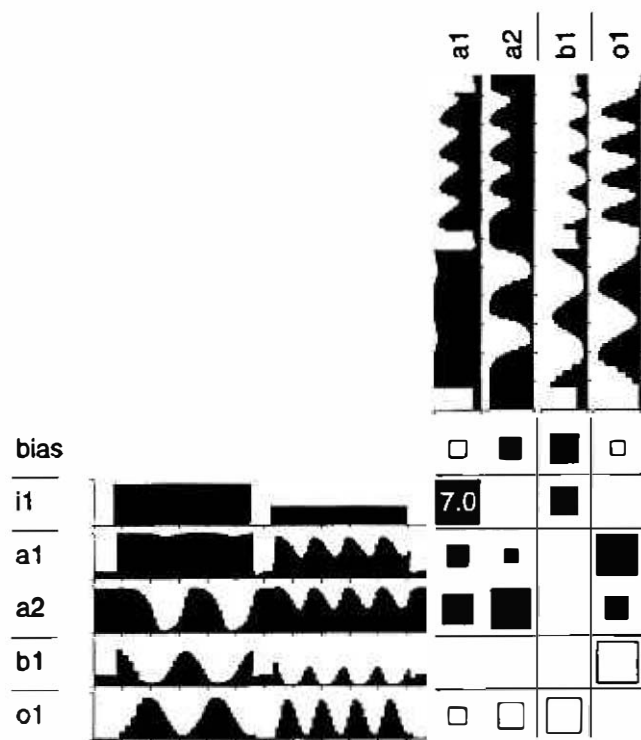


Network Trained to Generate Gated Sinusoidal Oscillation

Figure 7.3 The weight matrix gives the strength of connections from the row to column units (scale at top; excitatory black). The units are represented by name and activation; from top to bottom: the bias (a constant maximal activation of 1, not illustrated), the input gate signal (IN), the excitatory hidden units (e's), inhibitory hidden units (i's) and output (OUT).

is proportional to the strength of the connection, except those that exceed the calibrated scale at the top, which are designated numerically. In this simulation the excitatory hidden units are connected to each other, to the inhibitory group, and reciprocally to the output cell; the inhibitory hidden units connect only to the output. The network was trained without teacher forcing and achieved an oscillatory period of 20 time bins.

The network in figure 7.3 has more units than required to generate the oscillations. Many units have similar activations and connections (e.g., e4 and e9), and other units have negligible activity (e.g., e1 and i6). To simplify the network and elucidate the neural mechanisms, it is



Reduced Network That Generates Oscillations at Different Frequencies, Depending on the Level of Tonic Input

Figure 7.4 After the network was trained on the two illustrated input levels (0.4 and 0.9), testing with intervening levels of input generated oscillations at intervening frequencies.

often useful to reduce the network to its minimal size (LeCun et al. 1990; Mozer and Smolensky, 1989). A network of many hidden units can be reduced by (1) combining units with similar responses and connections into one equivalent unit and (2) eliminating units with negligible activation or weak connections. This can be done manually or by incorporating an automatic weight decay algorithm. Such reduced networks can then be retrained to perform the same operation. Figure 7.4 illustrates a reduced version of a network trained by my co-workers A. Pralle and L. Shupe to generate oscillations at two different frequencies, depending on the level of input. This network was trained to oscillate at a low frequency for an input of 0.9 and a high frequency for an input of 0.4. When presented with input levels between the training levels, this network also generates intervening frequencies. The interactions generating the oscillations can be analyzed by determining how the activation of each unit is determined by its inputs.

These simulations suggest that recurrent networks of sigmoidal units are quite robust in generating oscillatory activity, even to the point of meeting a variety of constraints in phase and period.

Reflex Responses of the Leech

Biological systems provide many examples of spatiotemporal transforms in reflex responses to localized stimuli. The leech withdrawal reflex has been modeled with dynamic recurrent networks by Lockery and Sejnowski (1990; 1992). The network inputs were stimuli applied to dorsal and ventral cutaneous receptors; the output signals were the

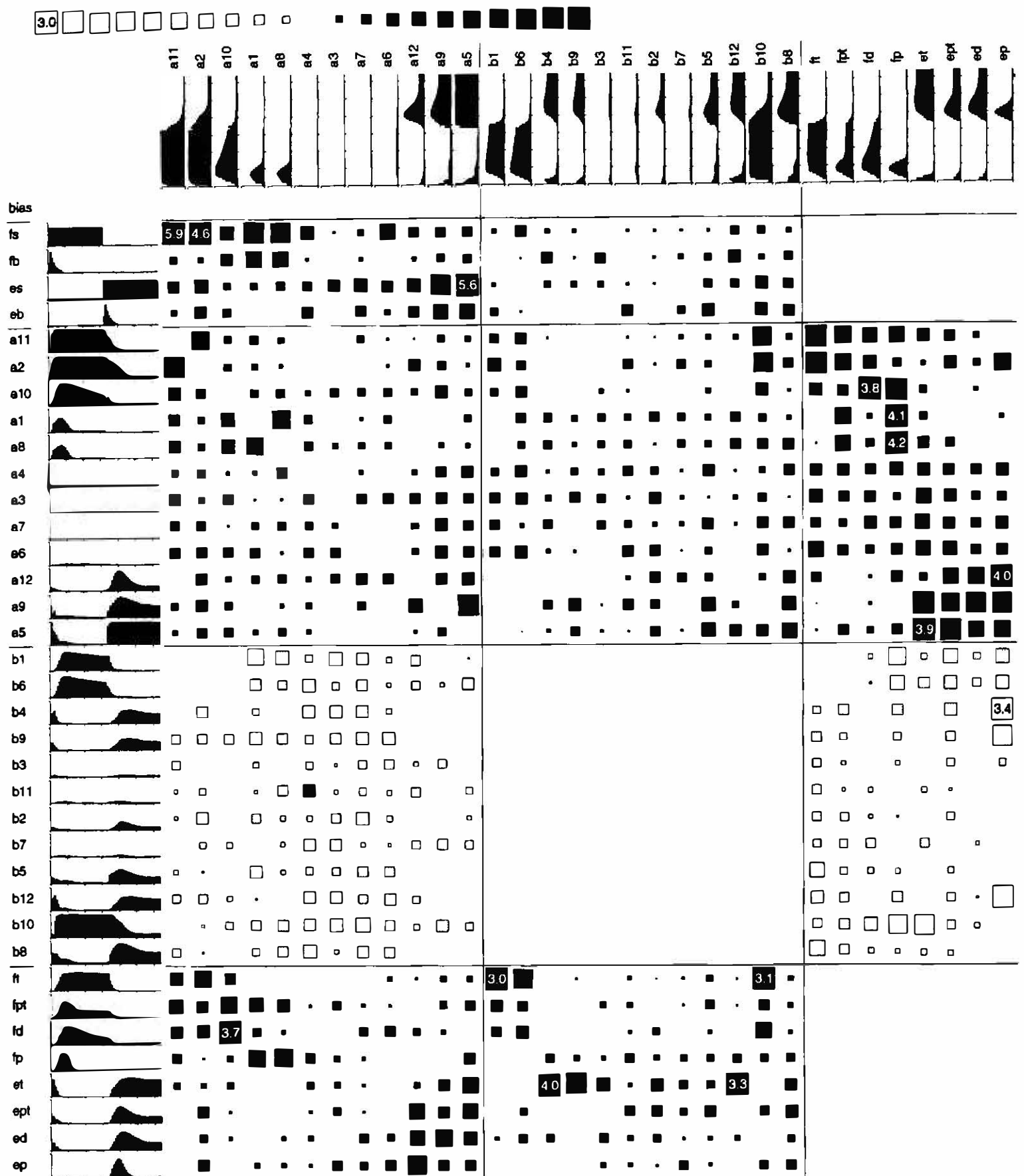
postsynaptic potentials evoked in representative motor neurons. These networks also incorporated electrical synapses and time constants which generated fast and slow components of synaptic potentials. The trained networks implemented distributed reflex pathways, as seen in the leech. Further details are provided in chapter 5.

Primate Step Tracking

The primary interest of my laboratory in dynamic recurrent networks derives from a desire to understand the neural circuitry controlling forelimb muscles of the primate. This will be discussed in some detail to illustrate the types of insights provided by network simulations. In monkeys performing a step-tracking task, previous experiments have documented the physiological discharge patterns and output connections of task-related neurons. The premotoneuronal (PreM) cells which affect muscle activity were identified in behaving monkeys by post-spike facilitation of target muscles in spike-triggered averages of electromyographic recordings. The response patterns of three groups of PreM cells—corticomotoneuronal (CM), rubromotoneuronal (RM), and dorsal root ganglion afferents—have been documented during a simple alternating flexion/extension task designed to relate their activity to changing and sustained force (Fetz et al. 1989; Flament et al. 1992). The discharge patterns observed in PreM cells, as well as in single motor units (MU) of agonist muscles (Palmer and Fetz 1985), fall into specific classes (Fetz et al. 1989). During a ramp-and-hold movement, all three groups include cells that show phasic-tonic discharge. The phasic component is related to the changing force and the tonic component is proportional to the amount of static force exerted. All groups also include tonic cells which show steady discharge throughout the hold period in proportion to the active force. Each group also has cells with unique firing properties. A large proportion of motor units show decrementing discharge, which decreases gradually through the static hold period. The RM population has cells that fire during both flexion and extension, and some cells that are unmodulated with the task and provide a constant bias.

To investigate the possible functional role of all these cells and to determine what other types of discharge patterns might be required to transform a step signal to the observed output of motor neurons, we derived dynamic networks that generated as outputs the average firing rates of motor units recorded in monkeys performing a step-tracking task.

Figure 7.5 shows a neural network that simulates the input-output transformation in the step-tracking task. The monkey sees a step change in target position, alternating between flexion and extension target zones. This is represented by sustained step inputs to the network for both flexion (fs) and extension (es). Since many visual system cells discharge



Dynamic Network Simulating Step-tracking Behavior Documented in Monkeys

Figure 7.5 The network transforms step and transient inputs to firing patterns of motor units at output. Unit activations during a flexion-extension cycle are shown at left and along the top. The rows represent, from top to bottom, the bias (which was eliminated for this simulation), the input signals (fs to eb), the excitatory hidden units (a's), inhibitory hidden units (b's) and the output flexor and extensor motor units (ft to ep). The target output patterns for each group are tonic (ft and et), phasic-tonic (fpt and ept), decrementing (fd and ed) and phasic (fp and ep). Hidden units were sorted in order of contribution to flexion vs. extension outputs. (Reprinted from Fetz and Shupe 1990, with permission.)

transiently, we also provided brief transient input at the onset of each target change (fb and eb). The network transforms these input signals to the observed response patterns of motor units at the output. The four types of motor unit patterns observed experimentally—tonic, phasic-tonic, decrementing, and phasic—are generated for both flexor and extensor movements (ft to ep). The intervening hidden units consist of twelve excitatory and twelve inhibitory neurons (a and b units, respectively).

Figure 7.5 illustrates the activation patterns and the connection matrix of all units after the network produces the eight different output patterns (2000 training iterations were sufficient). The discharge pattern of each unit is shown along the left, next to the row of output weights of that unit, and is shown again at the top of the column of weights representing the input connections to that unit. Self-recurrent connections (corresponding to weights on the diagonal) were excluded. To better visualize the relationships between units, the hidden units were sorted in order of the strength of their contribution to the phasic-tonic output units. The sorting algorithm used the product of activation and weight to the flexion phasic-tonic motor unit minus activation times weight to the extension phasic-tonic unit. Thus, the first hidden unit (a11) makes the largest relative contribution to the flexion phasic-tonic output unit (fpt). This hidden unit developed the strongest weights to the flexion tonic output unit (ft), but was also connected to the other flexor units. Such divergent connections to different motor units, as well as to synergist muscles are representative of CM cells (Fetz et al. 1989).

The activation patterns of the hidden units show several interesting features. The discharge patterns in the hidden units involve some recognizable variants of the output patterns, including tonic, phasic, phasic-tonic, and decrementing patterns. Secondly, although the activation profiles of the target motor units are identical for the flexion and the extension groups, the network solution involves a different assignment of hidden units devoted to each. There are more excitatory and fewer inhibitory hidden unit activations related to flexion than to extension, yet they produce essentially identical output effects. In general, different network simulations with the same architecture and trained on the same task, but starting with different initial weights usually converged on different solutions. (See chapter 3 for a relation of this property to biological variability.)

Another interesting feature of the weight matrix is the preferentially strong connections within the sets of units with sustained and transient activity. Thus, the first two flexor hidden units exhibiting tonic activity (a11 and a2) are strongly interconnected with each other and receive potent input from the flexion step and connect strongly to the tonic output unit. Similarly, the transient input (fb) is most strongly connected to the phasic hidden units (a1 and a8), which are strongly inter-

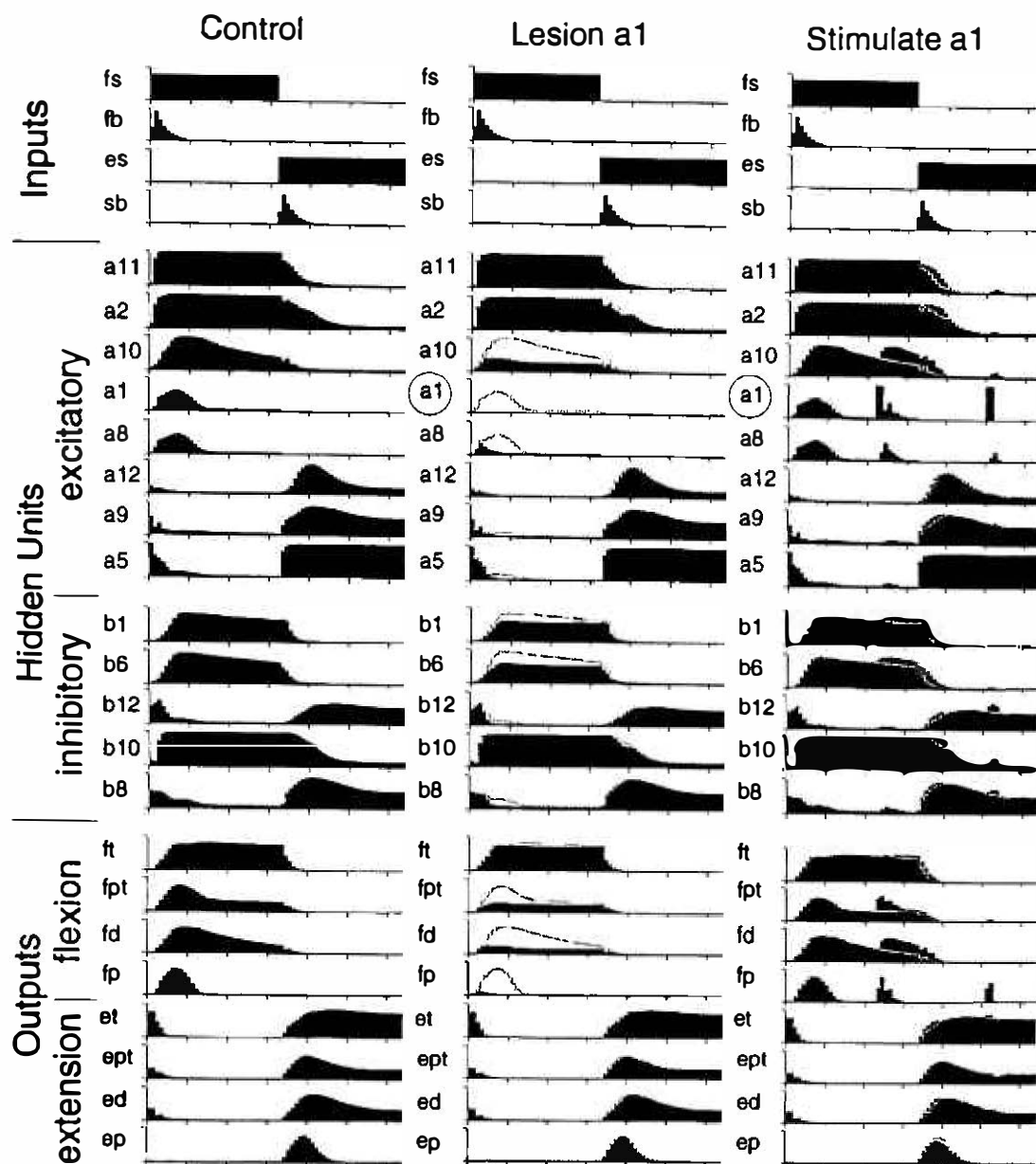
connected with each other and which also have strong reciprocal connections with the phasic output unit (fp).

Despite this tendency for sustained and transient signals to propagate through segregated pathways, most cells receive a mixture of phasic and tonic input signals. For example, the phasic flexor motor unit (fp) receives input not only from the phasic hidden units (a1 and a8), but also from cells with sustained activity (a11 and a10); indeed, some of the phasic activity of fp is derived from the difference between excitatory and inhibitory tonic cells with different onset times (e.g., a11 and b1). Such a subtraction of a delayed tonic pattern is used to produce phasic output in network simulations that receive only step inputs (Fetz et al. 1990).

We expected the network to develop reciprocal inhibition between the flexion and extension groups. Although a few inhibitory units activated during one phase of movement (e.g., extension) do affect primarily reciprocally activated motor units (e.g., b8 and b5), just as many units inhibit coactivated motor units (e.g., b12). Interestingly, the major drive on the tonic inhibitory units with connections to coactivated motor units is derived from recurrent feedback from the motor units (e.g., et to b4 and b9, which in turn inhibit ep). This recurrent feedback makes these inhibitory hidden units more analogous to the Renshaw cells of the spinal cord than to the Ia inhibitory neurons. Another group of inhibitory hidden units develop connections to both flexor and extensor motor units (b1 and b10), a pattern that has no known correlate in physiological studies.

The function of hidden units in the network can be tested by making selective *lesions*—by eliminating the activation of particular units and analyzing the behavior of the remaining network. Figure 7.6 illustrates the effects of removing a phasic hidden unit. The first column shows the control activations of representative units in the intact network. Removing the contribution of phasic unit a1 (middle column) eliminates the phasic activity of the output flexor units (fpt, and fp), as well as eliminating activity of the other phasic hidden unit (a8). It also reduces the sustained activity of the decrementing units (a10 and fd). These selective lesions have substantial effects because this network has relatively few units carrying a particular pattern, and these units are strongly interconnected.

The output effects of a given unit can also be tested by delivering a simulated *stimulus* and analyzing the propagated network response. The right column in figure 7.6 shows the effect of brief activation pulses delivered to a1 during the flexion and extension phase of the activity. The second pulse, during extension, evokes a response in the phasic output unit as well as in some of the active inhibitory hidden units. In contrast, the first pulse, delivered during the flexion phase, evokes a stronger response in the phasic output unit, as well as more prolonged



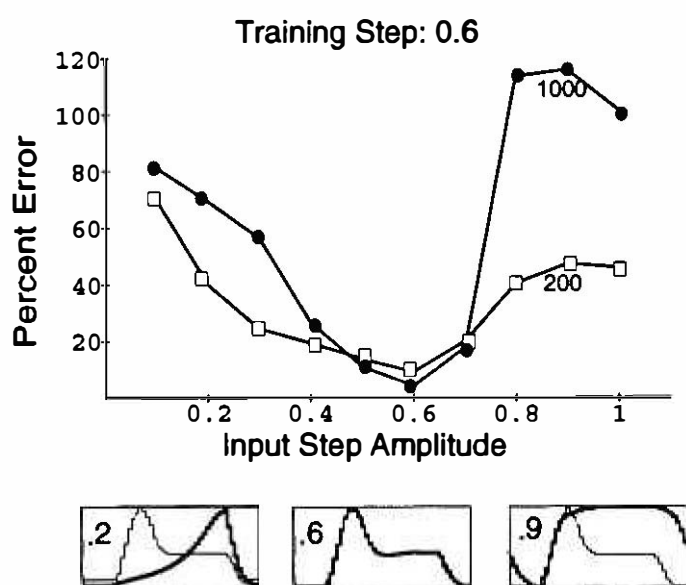
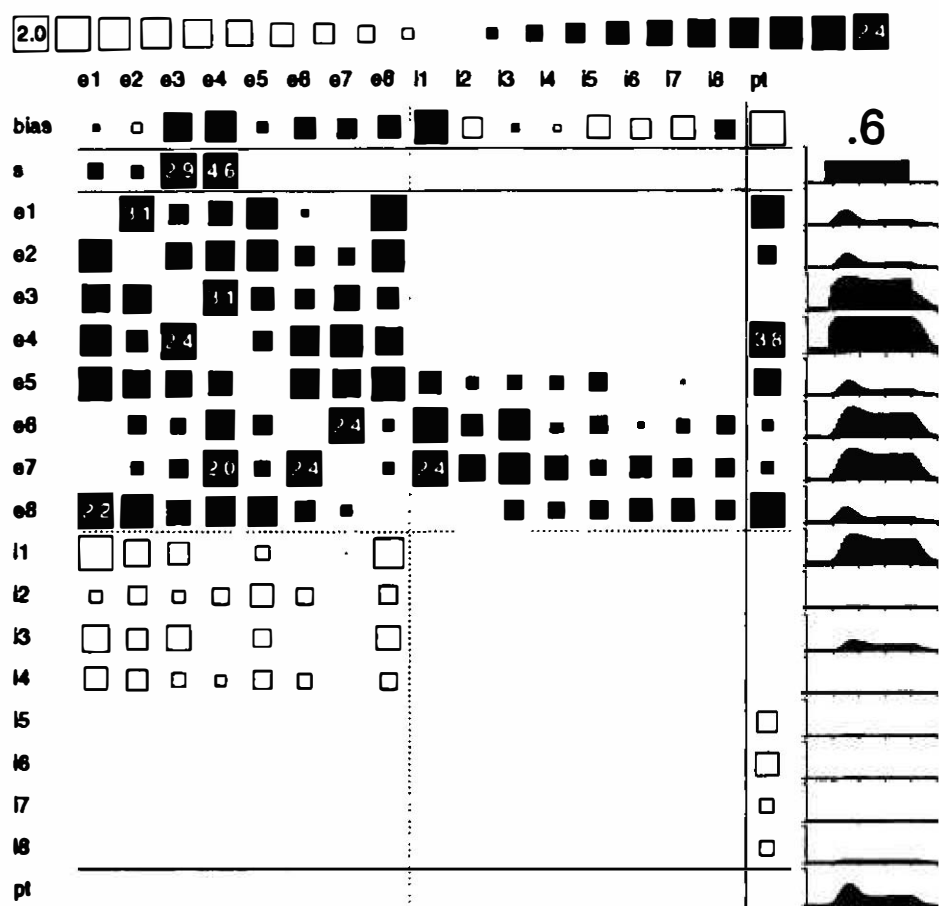
Effect of Lesioning and Stimulating a Hidden Unit in the Network of Figure 7.5

Figure 7.6 The activations of representative units are shown for the intact network (control, left column); after lesioning the phasic excitatory hidden unit a1 (middle column); and after stimulating the phasic excitatory hidden unit a1 (right column). The modified activations are shown in black, superimposed on a stippled profile of the control activations. (Reprinted from Fetz and Shupe 1990, with permission.)

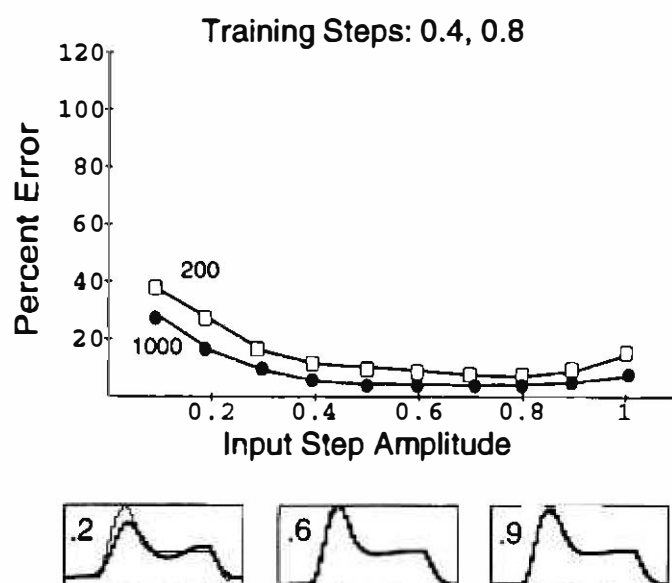
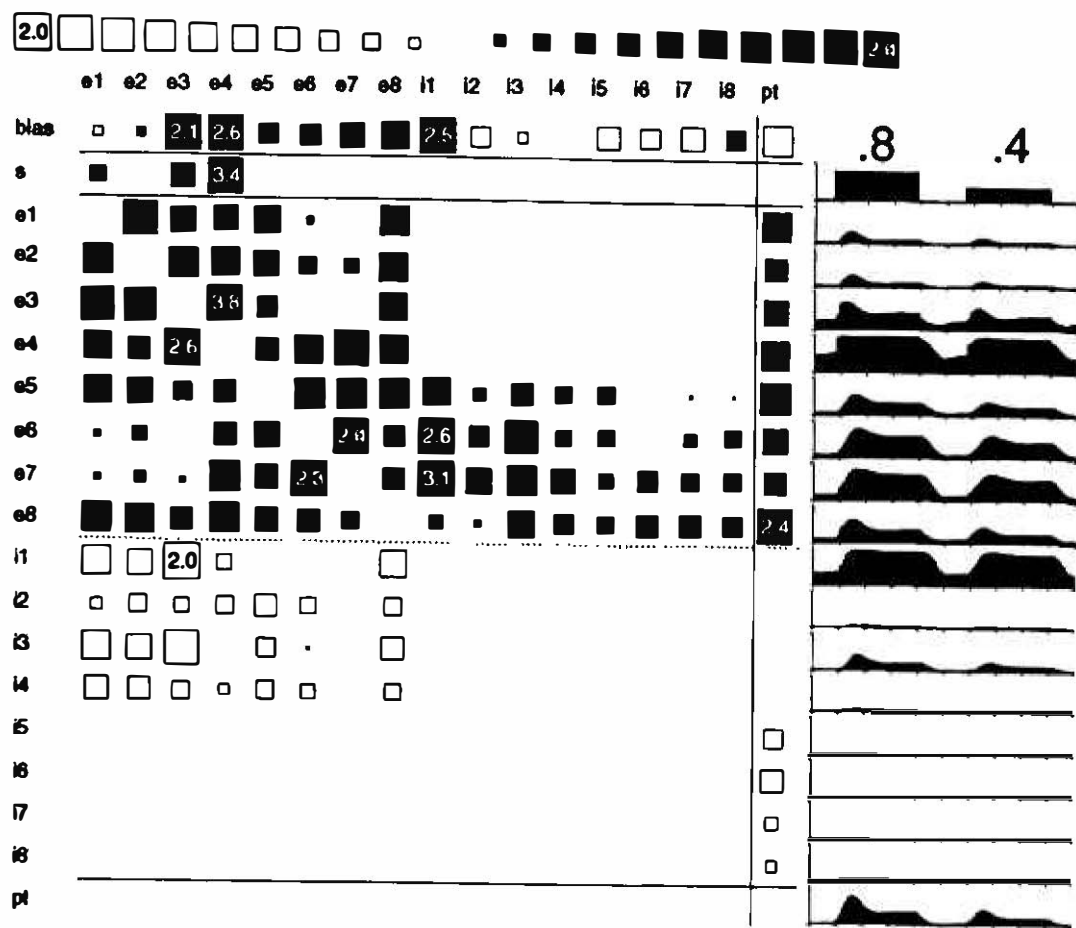
responses in the phasic-tonic and decrementing units (a10, fd), suggesting that the decrementing units carry a leaky integral of the phasic burst.

The fact that evoked responses were larger in certain phases of the movement than in others is clearly related to the gating function of the activation of the intervening units. Such modulation of evoked responses during cyclic motor activity has been well documented in physiological experiments. Plotting the evoked response as a function of time in the flexion-extension cycle revealed modulations that were not simply proportional to the activation of the stimulated or target cell, but were more complex functions of the activations and connections of the network (Fetz and Shupe 1990).

These networks were trained to generate the appropriate output patterns in response to specific and repeated inputs. We also investigated a network’s ability to *generalize* across different stimulus-response



magnitudes—to generate output activation patterns proportional to the input step. In the step-tracking task, monkeys learned to generate force levels proportional to the size of the target step, and did so by proportionally scaling the discharge levels of the task-related neurons. For example, the discharge pattern of phasic-tonic motor units was proportional to the active force levels (Palmer and Fetz 1985). Such linearity was tested in a simple network that transforms a step input to a phasic-tonic output (figure 7.7). If a network was trained on only one force level and tested with inputs representing other force levels it typically generated outputs that deviated drastically from a proportional phasic-tonic pattern (figure 7.7, left). Going from 200 to 1000 training iterations produced a slight improvement in accuracy at the training level (0.6) but resulted in greater deviations at other levels. We obtained a network that generates a phasic-tonic output pattern in proportion to the input



Magnitude Scaling in Neural Networks Transforming Step Input to Phasic-tonic Output

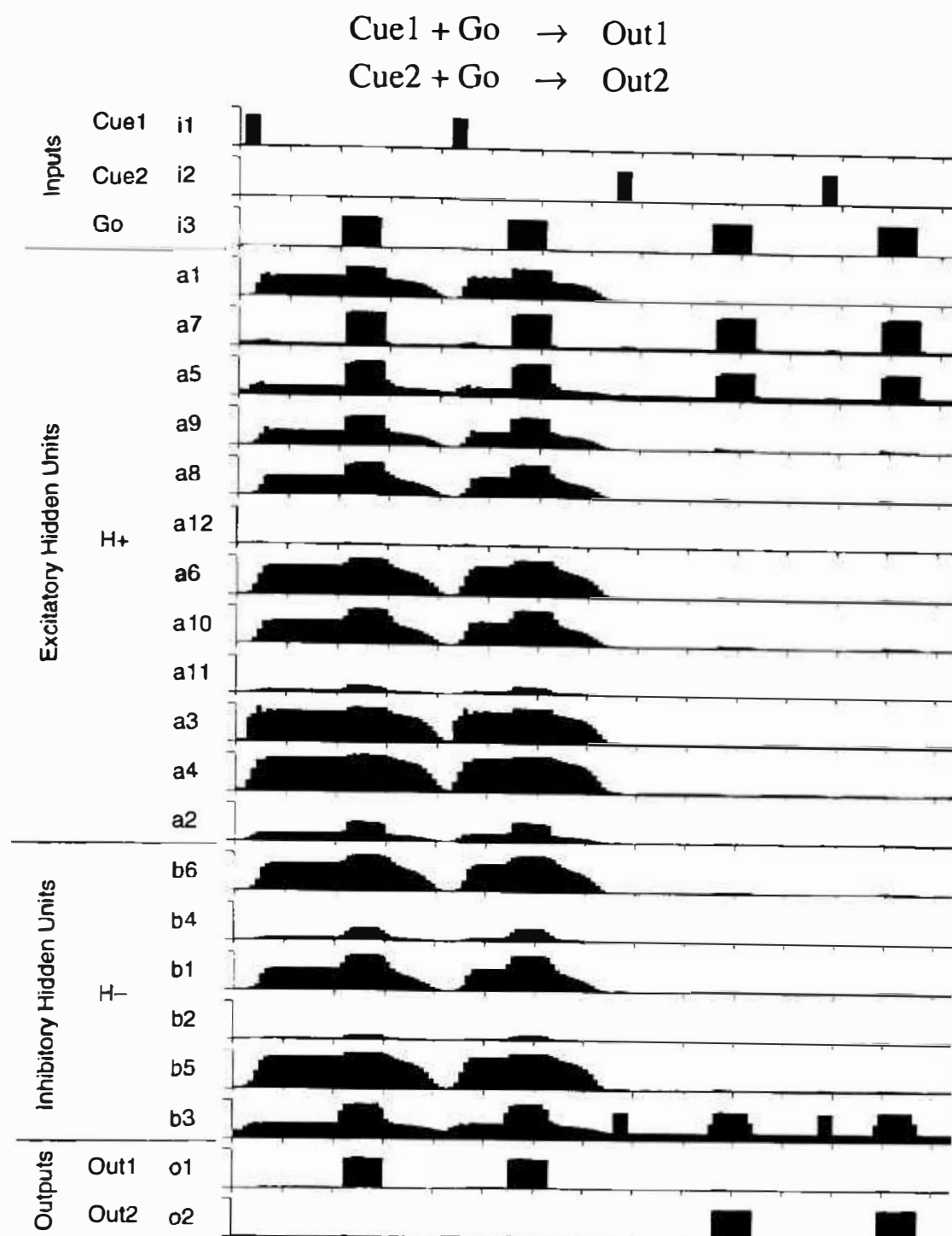
Figure 7.7 (Left) Network trained on one step size (0.6). (Right) Network trained from same starting point on two step sizes (0.4 and 0.8). Bottom graphs give percent error between the actual output and a phasic-tonic pattern proportional to the step input, after 200 training iterations (white boxes) and 1000 iterations (black circles). Samples below show normalized outputs (dark lines) superimposed on the phasic-tonic pattern (light lines) for input step sizes 0.2, 0.6 and 0.9, using the networks after 1000 iterations. (Reprinted from Fetz et al. 1990, with permission.)

step by simultaneous training at two step levels (0.8 and 0.4) (figure 7.7, right). When tested with intermediate input steps, the second network still generated proportional outputs. However, when tested with steps outside the range straddled by the training levels, the output still deviates slightly from a proportional response. Interestingly, the network that generalized the transform had a connection weight matrix similar to the network trained at a single level (both were trained from the same initial state); however, the network which generalized had a more equitable distribution of weights from the excitatory hidden units to the output (compare left and right networks in figure 7.7).

These simulations are obviously too simplistic to be taken as realistic models of the primate motor system. Nevertheless, even these highly abstracted networks reflect many of the features of the biological system in the monkey. For example, many different combinations of hidden units contribute to the same output. A phasic output pattern, for instance, has been generated by phasic hidden units, tonic excitatory activity minus delayed tonic inhibitory activity, and phasic-tonic excitatory input minus inhibitory bias. Furthermore, some complex activity patterns seen in premotor neurons of monkeys, such as bidirectional responses of RM cells, also appear in the networks. Even certain apparently paradoxical relations seen in monkeys, such as cortical units that covary with muscles which they inhibit, appear in networks and make contributions that are understandable in terms of other units. Thus, the networks have been useful in elucidating the function of many puzzling features of biological networks.

Short-term Memory Tasks

Neural mechanisms of short-term memory have been investigated in many experiments by recording cortical cell activity in animals performing instructed delay tasks. One variant of these tasks begins with an instruction cue that the monkey has to remember during a delay period. At the end of the delay a *go* stimulus signals the time to execute the appropriate response indicated by the cue. A simulation of this paradigm in a dynamic recurrent network is illustrated in figure 7.8. The network receives three inputs, consisting of the two separate cues and the *go* signal; the network has two outputs, corresponding to the response appropriate for each cue. Thus, following either cue 1 or 2, the *go* signal will generate activation of the corresponding output unit, 1 or 2. Figure 7.8 illustrates activations of the units in a trained network for several trials involving different delays. The connectivity matrix shows that the hidden units are configured to form a flip-flop that routes the *go* signal to the appropriate output. This network simulates the operation of short-term memory during the delay; some hidden unit activations (e.g., b3) carry mixed representations of cue and movement signals, which has been seen in premotor and prefrontal cortex neurons.

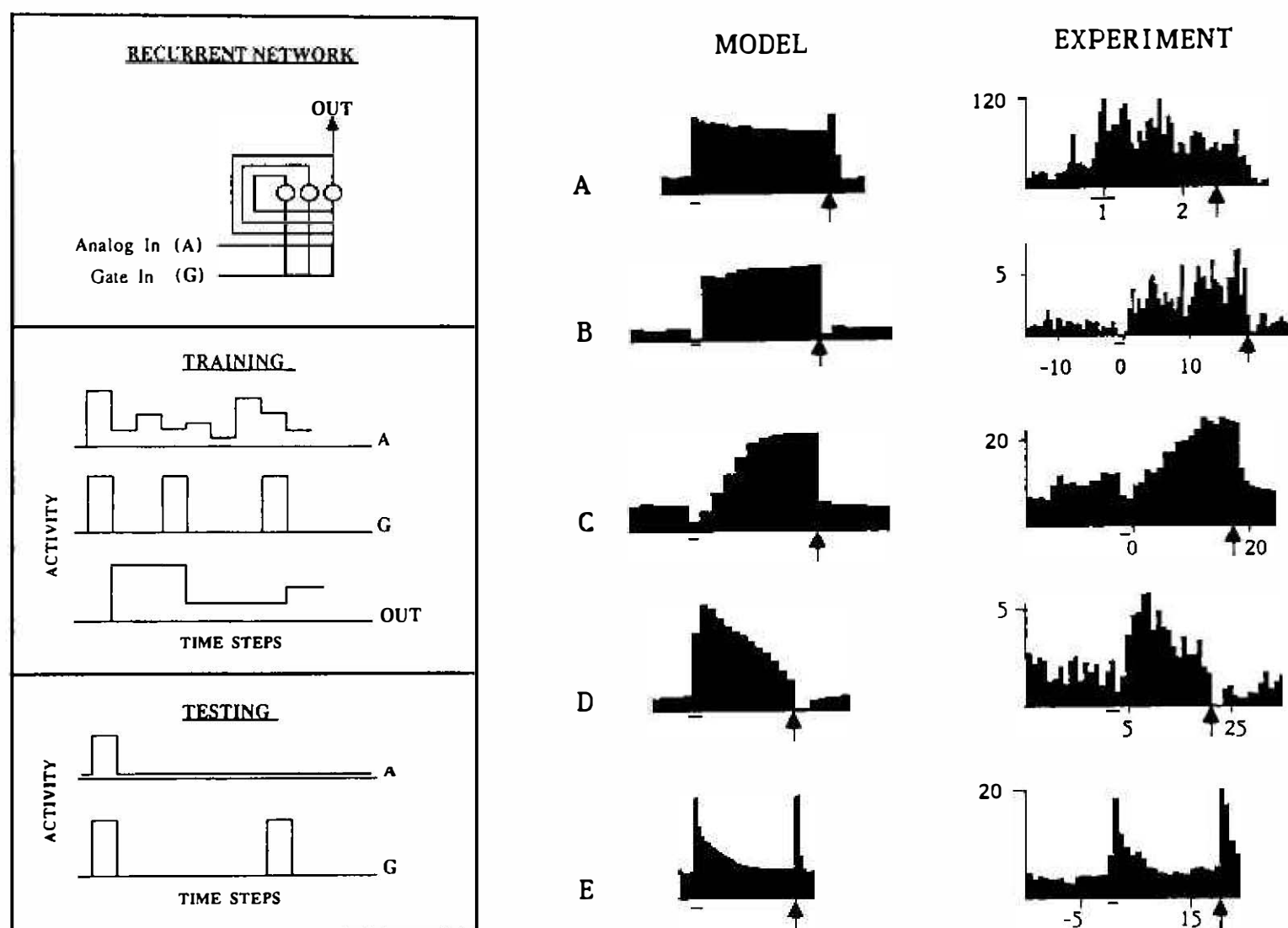


Activity Patterns in Network Simulating the Instructed Delay Task

Figure 7.8 Input pulses on inputs i1 and i2 are cues for the network to generate outputs o1 and o2, respectively, when the go signal i3 is presented after a delay.

In this simulation the output pulse is a replica of the go signal; this output pulse can be readily shaped to generate activation profiles resembling muscle activity by further training, or by cascading a second network that transforms the output pulse (as in figure 7.7).

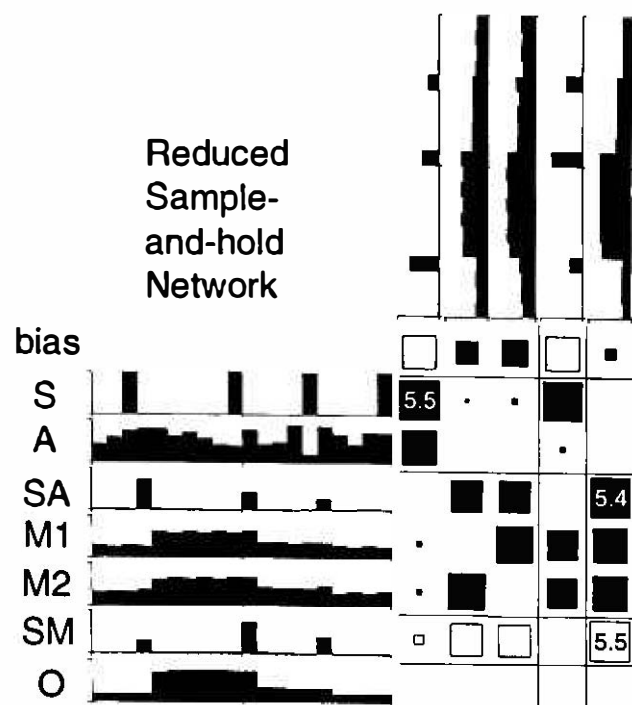
Another type of instructed delay task involves the requirement to remember the value of a particular stimulus. Williams and Zipser (1989a) trained recurrent networks to perform delayed nonmatch to sample task, in which the sampled variable to be remembered had binary values. More generally, the remembered variable can range over a continuum of values. Zipser (1990, 1991) has trained recurrent networks to simulate short-term memory of an analog value during the delay; in essence the network implements a sample-and-hold function (figure 7.9, left). The network has two inputs: an analog signal representing the stimulus



Sample-and-hold Delayed Response Task of Zipser (1990, 1991)

Figure 7.9 (Left) Simplified representation of input-output operation of the model, and waveforms used for training and testing. (Right) Comparison of activations of hidden units of the model and neural activity recorded in cerebral cortex of monkeys performing conditional delay tasks. The brief bar and arrow under the time axes represent the initial and final gate for the model, and the cues at the beginning and end of the delay period for the experimental data. The experimental histograms represent: (A) a neuron from posterior parietal area LIP during a delayed saccade task (Gnadt and Andersen 1988); (B) an inferotemporal neuron during a visual delayed match-to-sample task (Fuster et al. 1985); (C and E) frontal neurons in the principal sulcus during a delayed match-to-sample task (Fuster 1984); (D) a frontal cortex neuron during a delay choice task (Quintana et al. 1989). (Reproduced with permission from Zipser 1990, 1991; see these references for further information.)

value to be remembered and a gate signal specifying the sample and output times. The network output is the value of the analog input at the time of the previous gate. During the delay the activity of many hidden units resembles the response patterns of cortical neurons recorded in monkeys performing comparable instructed delay tasks (figure 7.9, right). The activity patterns of hidden units, like those of cortical neurons, fall into three main classes: activation profiles proportional to the remembered analog value, often with a decay; activation during the gate signal; and some combination of the two. The network simulations allow the function of each of these patterns observed in the animal to be interpreted in terms of its role in the task.



Reduced Network Performing a Sample-and-hold Function, Simulating Short-term Memory

Figure 7.10 A larger network with weight constraints was originally trained on the task, then reduced to the essential minimum by eliminating redundant and unnecessary hidden units. The two inputs are the sample signal (S) and a random analog value A; the output (O) is the sustained value of the last sampled analog value.

My laboratory has also investigated such short-term memory networks in order to further analyze their operation. To elucidate the underlying computational algorithm we have found it useful to constrain units to have either excitatory or inhibitory output weights (not both), and to reduce the network to the minimal essential network that performs the same function. A reduced network performing the sample-and-hold function with sign constraints on the weights is illustrated in figure 7.10. It consists of three excitatory and one inhibitory unit. The two inputs are the sample gate signal (S) and the analog variable A; the output (O) is the value of A at the last sample gate. This reduced version reveals an elegant computational algorithm that exploits the nonlinear sigmoidal input-output function of the units. The first excitatory unit (SA) carries a transient signal proportional to the value of A at the time of the gate. This signal is derived by clipping the sum of the analog and gating inputs with a negative bias, as shown by the input weights to SA in the first column. This input sample is then held in memory by two excitatory units (M1 and M2), which maintain their activity by reciprocal connections and which feed their summed activity to the output. The inhibitory unit (SM) carries a transient signal proportional to the previous value of A. Its value is derived from a clipped sum of the gate S and the previous values held in M1 and M2. The function of SM is to subtract the previously held value from the integrating hidden units and from the output. Thus, the network reveals an elegant use of nonlinearity and integration to yield the appropriate remembered value.

It seems likely that networks with more units implement a comparable algorithm in a distributed manner.

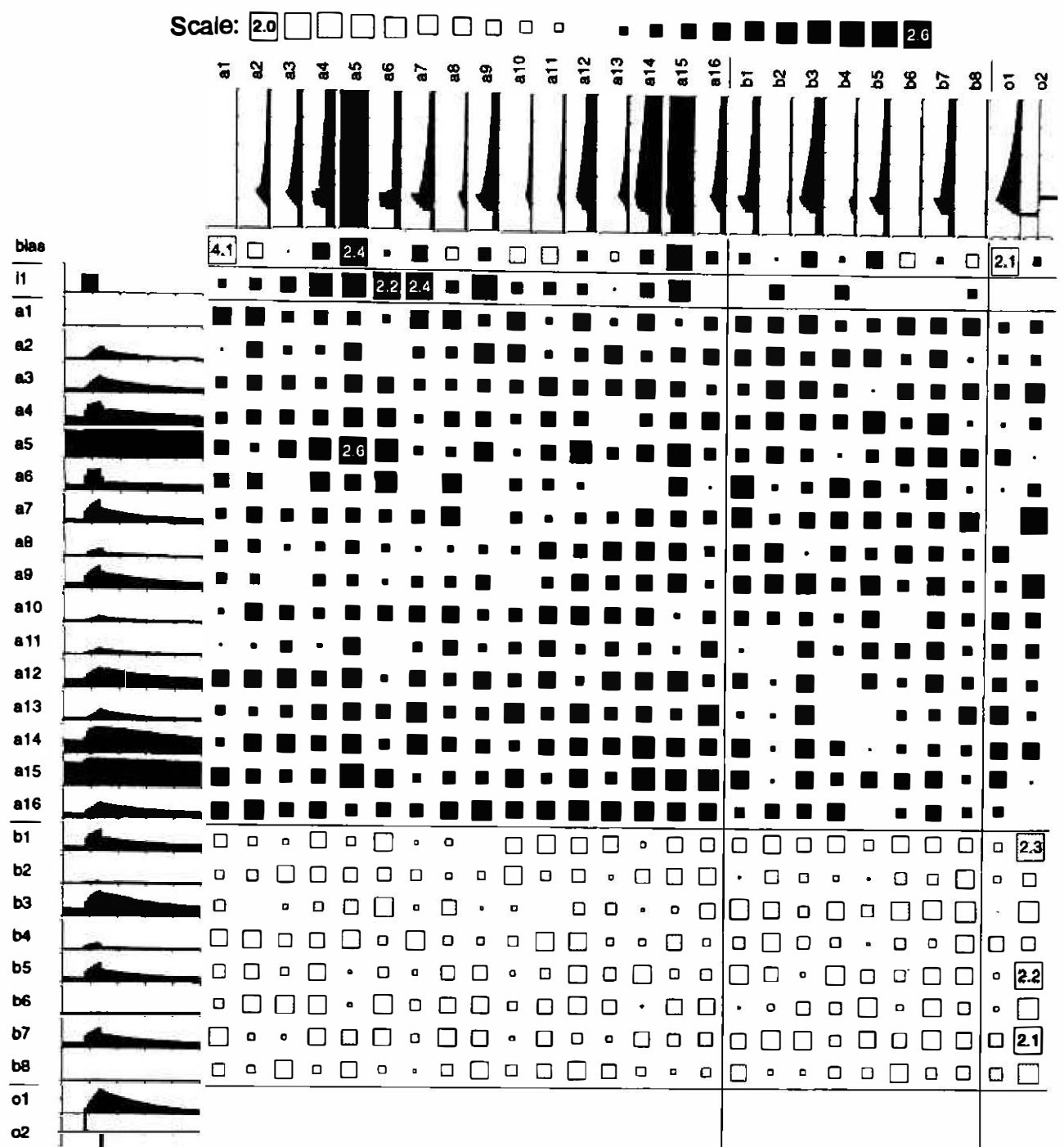
Neural Integration and Differentiation

Biological systems provide many examples of networks that integrate and differentiate neural activity. Munro et al. (1991) trained dynamic recurrent networks with backpropagation to generate the integral and differential of arbitrary input patterns. The connectivity of these networks was not designed to model any particular biological system; instead, our purpose was to investigate the mechanisms of neural computation of analytical functions and test the possibility of implementing multiplexed functions. Recurrent networks with architectures similar to that in figure 7.2 were trained with inputs of quasi-sinusoidal waveforms to produce either the leaky integral or the differential of the input or both. The nonrepeating input signal was a sinusoidal waveform whose frequency was varied up and down in random steps. The desired target output waveforms were calculated as the differential or the leaky integral delayed by two timesteps. Varying the frequency of the training signal insured that the network performed accurately over a broad range of frequencies. The trained networks compute these transforms for any arbitrary inputs.

By eliminating unnecessary units and combining redundant units and retraining the remaining network, it was possible to derive reduced networks that performed integration or differentiation or both. The computational algorithm was often clearer in the reduced networks. (However, smaller networks could not be trained from random starting weights to perform these functions.) The differentiating networks implemented a strategy of delayed subtraction of the input, and their hidden units often carried signals resembling the input. The integrator networks used recurrent connections between the excitatory hidden units, whose activation all resembled the output—i.e., the integral of the input.

Networks could also be trained to produce the integral and differential simultaneously on two separate outputs, as in figure 7.11 (Munro et al. 1991). These networks combined the strategies used by the pure differentiator and integrator in a distributed and overlapping manner. In general, a large number of hidden units were actively involved in the multiplexed computation and many of these contributed significantly to both outputs (compare vertical column of input weights to o1 and o2). The recurrent connectivity of the excitatory units is similar to that obtained for a simple integrator network. The activity patterns of most hidden units (both excitatory and inhibitory) carry a component that resembles the leaky integral of the input.

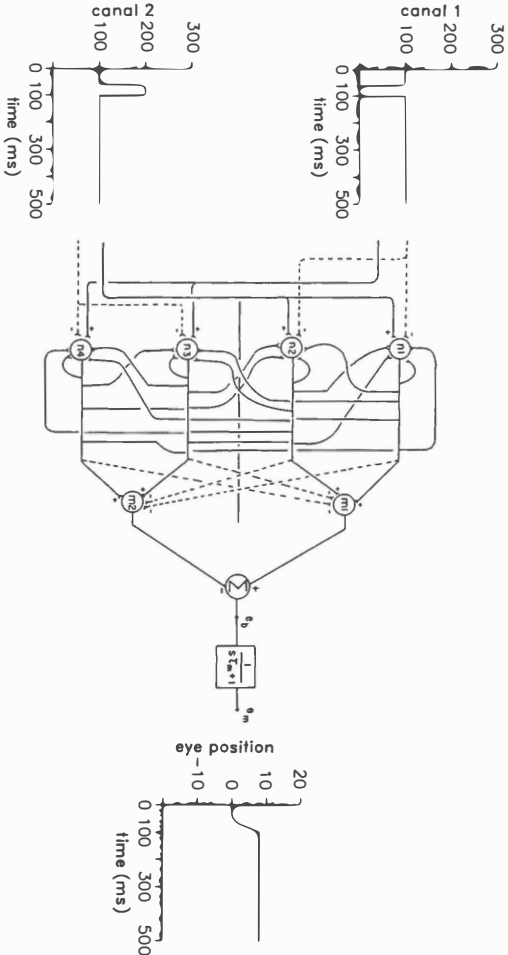
In biological motor systems, neural integrators have been postulated to transform transient commands into sustained activity, and to mediate the vestibulo-ocular reflex (VOR). Arnold and Robinson (1991) modeled



Dynamic Recurrent Network Generating Integral and Differential of the Input

Figure 7.11 Activity of a dynamic recurrent network trained with nonrepeating input waveforms to generate both the leaky integral (o1) and the differential (o2) of the input. The illustrated input waveform was a pulse, to show the activations of the hidden and output units.

the vestibulo-ocular integrator with a recurrent network whose connections resembled those of the vestibulo-ocular system (figure 7.12). The two input signals were representations of the reciprocal responses of opposed vestibular afferents to head movement; these connected in push-pull fashion to four interneurons which were interconnected to each other and to motor neurons. Since the vestibular afferents carry tonic activity in the absence of head movement, the integrator had to be configured so as to integrate only deviations from baseline, but not the baseline activity itself. The authors used units with an exponential decay and a rectifying input-output characteristic. To train units with this nondifferentiable function they used a gradient descent method that involved tweaking the weights individually by a small amount, deter-



Connectivity of Dynamic Network Simulating Vestibulo-ocular System

Figure 7.12. Two opponent vestibular canal inputs are connected to four recurrently connected interneurons, which in turn connect to antagonist oculomotor neurons. Connections of units can be both excitatory (solid lines) and inhibitory (dashed lines). Eye position is derived from the difference of the motor neurons' outputs. (Reproduced from Arnold and Robinson 1991, with permission.)

mining the effect on the error, and using this operationally derived estimate of the error gradient to update all weights. The resulting network is shown in figure 7.12 with the activity profiles of vestibular canal inputs and resultant eye movements. This network allowed units to have both excitatory and inhibitory output connections. Integration was performed through positive recurrent connections between the interneurons. Removing a hidden unit in a trained network reduced the time constant of integration, but the network could be readily retrained. The network could also mimic more complex responses, such as postsaccadic drift.

Anastasio (1991a, b) trained dynamic networks of sigmoidal units with backpropagation to simulate velocity storage in the vestibulo-ocular system. A three-layer network model was trained to produce compensatory long-duration activity of oculomotoneurons in response to brief head velocity inputs. The network learned to produce the prolonged response, termed velocity storage, by developing lateral inhibitory interactions between the interneurons. These interneurons exhibited many of the response properties of VOR interneurons in the vestibular nucleus—i.e., low baseline firing rates, long time constants, and rectified and skewed responses. The model indicates that these properties are related and are the result of lateral inhibition. The model further replicated the consequences of vestibular lesions.

Lisberger and Sejnowski (1992a) used dynamic networks to investigate mechanisms of learning in the vestibulo-ocular system (see also the extensive discussion in Churchland and Sejnowski 1992). The network was constructed to include many anatomical and physiological constraints, including pathways through the cerebellar flocculus, with appropriate delays. The two inputs to the network, head velocity and target velocity, were converted to a single output: eye velocity. The network was initially trained to simulate three behaviors: smooth pursuit of a moving visual target, the VOR to head movement, and suppression of the VOR (when head and target move together). Then the network was required to change the gain of the VOR, as occurs after wearing magnifying or minifying goggles, and also to maintain accurate smooth pursuit visual tracking. These requirements led to changes in the weights of connections at two specific sites: the vestibular input to the flocculus and to the brainstem neurons controlling oculomotoneurons. A change in the balance of input from phasic-tonic and tonic vestibular efferents contributed to a change in the gain of the VOR, because the transient signal was integrated by a positive feedback loop (Lisberger and Sejnowski 1992a, b). This study exemplifies the insights gained from a biologically constrained dynamic model that can incorporate the time course of neural activity observed under different behavioral conditions, and shows the power of such simulations to reveal novel network mechanisms.

FUTURE DIRECTIONS

These examples illustrate the use of dynamic recurrent networks to simulate sensorimotor behavior and to elucidate the underlying neural computations. To better understand the operations of biological networks, future development of these techniques can be directed toward incorporating more realistic neural properties into the units. The sigmoidal input-output function of the units is a mathematically convenient but artificial approximation to the response of biological neurons. One important property of physiological neurons is that their activity depends not only on immediate inputs, but also on their prior activity. History-dependent unit activity such as adaptation and postinhibitory rebound arises from the kinetics of time- and voltage-dependent ionic channels (Schwindt 1992; Spain et al. 1991; see also chapter 2). These properties have been amply documented for many neuronal types (reviewed by Llinás 1988). The activation kinetics of ionic channels make the neuron's activity depend on its past behavior as well as on its instantaneous input. These intrinsic modulations may be quite significant for the network solutions. For example, in monkeys performing the step-tracking task, the decrementing motor units may simply be exhibiting an adapting response to a tonic input. Our simulations to date using sigmoidal units were forced to generate this decrementing pattern by creating appropriate simultaneous input from the hidden units. Simulations with adapting units could generate both the decrementing and tonic motor unit responses to the same tonic inputs. The basic property of history-dependent activation can be modeled to a first approximation by passing the summed synaptic input through a leaky integrator in parallel with a direct gain element (Schwindt 1992). This has been accomplished by incorporating a passive decay in many simulations (Arnold and Robinson 1991; Lisberger and Sejnowski 1992a; Lockery and Sejnowski 1992). As physiological information becomes more detailed it will be possible to incorporate more specific types of time dependencies into network units.

A second feature of real neurons that may be quite significant for information processing is *discrete spiking*. The sigmoidal units transform a continuous function representing mean firing rate; in this sense, the units may be more properly interpreted as representing the mean activity of a large population of asynchronously spiking neurons. Real neurons integrate synaptic input to threshold and fire discrete spike trains, which may have functional consequences. For example, synchronous activity in spiking neurons can enhance the effectiveness of transmission of signals (Kenyon et al. 1990). Unfortunately, networks with spiking neurons cannot be trained using backpropagation. Other training methods, reviewed below, could be used to change the network weights. However, the training is much slower because many iterations

must be averaged before the results of each weight change can be assessed.

Although backpropagation is very effective in searching out solutions in weight space, it is fraught with uncertainty about finding the best solutions, and can only be used with units whose input-output function f is differentiable. Other training strategies can be used to find network solutions with more realistic units. For example, *random search strategies* (Baba 1989; Choi et al. 1991) sample a number of randomly selected points distributed in weight space and then pursue the best solutions with random perturbations of the weights. This strategy is computationally simpler and faster *per step* than backpropagation, and can be applied for networks of units with nondifferentiable response functions.

Another important reinforcement strategy is A_{R-P} training (for *associative reward-penalty*), in which all the weights are modified in proportion to the output error and in proportion to local presynaptic and postsynaptic activations (Barto and Jordan 1987). This modified Hebb rule has greater biological plausibility, since reward and punishment would generally have global effects rather than specifically changing each weight in the direction of reducing error. In direct comparisons with backpropagation, A_{R-P} training converged to similar network solutions, albeit more slowly (Mazzoni et al. 1991a, b). These two training algorithms are both applicable to neurons which have biologically plausible and nondifferentiable characteristics.

CONCLUSIONS

The unique insights provided by neural network simulations ensure their continued use in elucidating the operations of neural systems. The basic limitation of conventional physiological and anatomical data is that they provide a very selective sample of a complex system, leaving a wide gap between particular glimpses of neural activity or anatomical structure, and the behavior of the overall system. This gap is usually bridged by intuitive inferences, often based on selective interpretation of the data (Fetz 1992). A more objective method of bridging the gap is with simulations that provide complete network models. These models can incorporate the observed responses of units and can help explain the functional meaning of neural patterns. Thus, systems neurophysiologists can profitably use a combination of unit recording techniques and neural modeling to elucidate the network mechanisms generating sensorimotor behavior. Unit recordings can provide important information on the activity of related neurons, but the network models can provide working examples of complete solutions to sensorimotor behavior. To the extent that models can incorporate anatomical and physiological constraints, they can provide plausible explanations of the neural mechanisms underlying behavior.

ACKNOWLEDGMENTS

I gratefully acknowledge the invaluable help of Mr. Larry Shupe, who wrote the software for our dynamic recurrent network simulator, and has helped explore many of its applications. Others who worked on various simulations described here include Edmund Munro, Soren Impey, Venky Murthy, and Arnd Pralle. This work was initially supported by the Office of Naval Research (contract number N00018-89-J-1240), and by NIH grants RR00166 and NS12542.

References

- Abraham, W.C. and Goddard, G.V. 1983. Asymmetric relationships between homosynaptic long-term potentiation and heterosynaptic long-term depression. *Nature* 305: 717–719.
- Abraham, W.C., Bliss, T.V.P., and Goddard, G.V. 1985. Heterosynaptic changes accompany long-term but not short-term potentiation of the perforant path in the anesthetized rat. *Journal of Physiology (London)* 363: 335–349.
- Abrams, T.W. and Kandel, E.R. 1988. Is contiguity detection in classical conditioning a system or a cellular property? Learning in *Aplysia* suggests a possible molecular site. *Trends in Neurosciences* 4: 128–135.
- Abrams, T.W., Karl K.A., and Kandel, E.R. 1991. Biochemical studies of stimulus convergence during classical conditioning in *Aplysia*: dual regulation of adenylate cyclase by Ca^{2+} / calmodulin and transmitter. *Journal of Neuroscience* 11: 2655–2665.
- Ackley, D.H., Hinton, G.E., and Sejnowski, T.J. 1985. A learning algorithm for Boltzmann machines. *Cognitive Science* 9: 147–169.
- Alonso, A., De Curtis, M., and Llinás, R.R. 1990. Postsynaptic Hebbian and non-Hebbian long-term potentiation of synaptic efficacy in the entorhinal cortex in slices and in the isolated adult guinea pig brain. *Proceedings of the National Academy of Sciences USA* 87: 9280–9284.
- Alonso-deFlorida, F., Morales, M.A., and Minzoni, A.A. 1991. Modulated long-term potentiation in the cat superior cervical ganglion *in vivo*. *Brain Research* 544: 203–210.
- Anastasio, T.J. 1991a. Neural network models of velocity storage in the horizontal vestibulo-ocular reflex. *Biological Cybernetics* 64: 187–196.
- Anastasio, T. J. 1991b. A recurrent neural network model of velocity storage in the vestibulo-ocular reflex. In: *Advances in Neural Information Processing Systems 3*, edited by R.P. Lippmann, J.E. Moody, and D.S. Touretzky. San Mateo, Calif.: Morgan Kaufmann Publishers, pp. 32–38.
- Anastasio, T.J. 1992. Simulating vestibular compensation using recurrent back-propagation. *Biological Cybernetics* 66: 389–397.
- Anastasio, T.J. and Robinson, D.A. 1990. Distributed parallel processing in the vertical vestibulo-ocular reflex: learning networks compared to tensor theory. *Biological Cybernetics* 63: 161–167.
- Anderson, J.A. and Rosenfeld, E., eds. 1988. *Neurocomputing: Foundations of Research*. Cambridge, Mass.: MIT Press.
- Anderson, J.A., Pellionisz, A., and Rosenfeld, E., eds. 1990. *Neurocomputing 2: Directions for Research*. Cambridge, Mass.: MIT Press.

- Angstadt, J.D. and Calabrese, R.L. 1991. Calcium currents and graded synaptic transmission between heart interneurons of the leech. *Journal of Neuroscience* 11: 746–759.
- Arnold, D.B. and Robinson, D.A. 1990. A neural network that learns to integrate oculomotor signals. *Proceedings of the International Joint Conference on Neural Networks-90* II: 693–696.
- Arnold, D.B. and Robinson, D.A. 1991. A learning network model of the neural integrator of the oculomotor system. *Biological Cybernetics* 64: 447–454.
- Artola, A. and Singer, W. 1990. The involvement of N-methyl-D-aspartate receptors in induction and maintenance of long-term potentiation in rat visual cortex. *European Journal of Neuroscience* 2: 254–269.
- Artola, A., Bröcher, S., and Singer, W. 1990. Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature* 347: 69–72.
- Aszódi, A. and Friedrich, P. 1987. Molecular modelling of associative learning. *Neuroscience* 22: 37–48.
- Aszódi, A., Müller, U., Friedrich, P., and Spatz, H.-C. 1991. Signal convergence of protein kinase A as a molecular correlate of learning. *Proceedings of the National Academy of Sciences USA* 88: 5832–5836.
- Baba, N. 1989. A new approach for finding the global minimum of error function of neural networks. *Neural Networks* 2: 367–373.
- Bachoo, M. and Polosa, C. 1991. Long-term potentiation of nicotinic transmission by a heterosynaptic mechanism in the stellate ganglion of the cat. *Journal of Neurophysiology* 65: 639–647.
- Bailey, C.H. and Chen, M. 1983. Morphological basis of long-term habituation and sensitization in *Aplysia*. *Science* 220: 91–93.
- Bailey, C.H. and Chen, M. 1988. Long-term memory in *Aplysia* modulates the total number of varicosities of single identified sensory neurons. *Proceedings of the National Academy of Sciences USA* 85: 2373–2377.
- Bailey, C.H., Chen, M., Keller, F., and Kandel, E.R. 1992. Serotonin-mediated endocytosis of apCAM: an early step of learning-related synaptic growth in *Aplysia*. *Science* 256: 645–649.
- Ballard, D.H. 1990. Modular learning in hierarchical neural networks. In: *Computational Neuroscience*, edited by E.L. Schwartz. Cambridge, Mass.: MIT Press, pp. 139–153.
- Barrionuevo, G., Schottler, F., and Lynch, G. 1980. The effects of repetitive low frequency stimulation on control and “potentiated” synaptic responses in the hippocampus. *Life Sciences* 27: 2385–2391.
- Barto, A.G. and Jordan, M.I. 1987. Gradient following without back-propagation in layered networks. *Proceedings of the IEEE First International Conference on Neural Networks* II: 629–636.
- Baudry, M. and Davis, J.L., eds. 1991. *Long-Term Potentiation: A Debate of Current Issues*. Cambridge, Mass.: MIT Press.
- Baux, G. and Tauc, L. 1987. Presynaptic actions of curare and atropine on quantal acetylcholine release at a central synapse of *Aplysia*. *Journal of Physiology (London)* 388: 665–680.
- Baux, G., Fossier, P., and Tauc, L. 1990. Histamine and FLRFamide regulate acetylcholine release at an identified synapse in *Aplysia* in opposite ways. *Journal of Physiology (London)* 429: 147–168.

- Baxter, D.A. and Byrne, J.H. 1990. Differential effects of cAMP and serotonin on membrane currents, action-potential duration, and excitability in somata of pleural sensory neurons of *Aplysia*. *Journal of Neurophysiology* 64: 978–990.
- Baxter, D.A., Bittner, G.D., and Brown, T.H. 1985. Quantal mechanism of long-term synaptic potentiation. *Proceedings of the National Academy of Sciences USA* 82: 5978–5982.
- Baxter, D.A., Buonomano, D.V., Raymond, J.L., Cook, D.G., Kuenzi, F.M., Carew, T.J., and Byrne, J.H. 1991. Empirically derived adaptive elements and networks simulate associative learning. In: *Neural Network Models of Conditioning and Action*, edited by M.L. Commons, S. Grossberg, and J.E.R. Staddon. Hillsdale, N.J.: Lawrence Erlbaum, p. 13–53.
- Bear, M.F., Cooper, L.N., and Ebner, F.F. 1987. A physiological basis for a theory of synaptic modification. *Science* 237: 42–48.
- Bear, M.F., Kleinschmidt, A., Gu, Q., and Singer, W. 1990. Disruption of experience-dependent synaptic modifications in striate cortex by infusion of an NMDA receptor antagonist. *Journal of Neuroscience* 10: 909–925.
- Becker, S. and Hinton, G.E. 1992. Self-organizing neural network that discovers surfaces in random-dot stereograms. *Nature* 355: 161–163.
- Beer, R.D. 1990. *Intelligence as Adaptive Behavior: An Experiment in Computational Neuroethology*. San Diego, Calif.: Academic Press.
- Beer, R.D. 1992. A dynamical systems perspective on autonomous agents. Technical Report, Department of Computer Engineering and Science, Case Western Reserve University, CES-92-11.
- Beer, R.D. and Gallagher, J.C. 1992. Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior* 1: 91–122.
- Beer, R.D., Chiel, H.J., and Sterling, L.S. 1989. Heterogeneous neural networks for adaptive behavior in dynamic environments. In: *Advances in Neural Information Processing Systems*, edited by D. S. Touretzky. San Mateo, Calif.: Morgan Kaufmann Publishers, pp. 577–585.
- Beer, R.D., Chiel, H.J., Quinn, R.D., Espenschied, K., and Larsson, P. 1992. A distributed neural network architecture for hexapod robot locomotion. *Neural Computation* 4: 356–365.
- Bekkers, J.M. and Stevens, C.F. 1989. Dual modes of excitatory synaptic transmission in the brain. In: *Proceedings of the First NIMH Conference*, edited by S. Zaleman and R. Scheller. Rockville, Maryland: Department of Health and Human Services, pp. 39–50.
- Bekkers, J.M. and Stevens, C.F. 1990a. Computational implications of NMDA receptor channels. *Cold Spring Harbor Symposia on Quantitative Biology* 55: 131–135.
- Bekkers, J.M. and Stevens, C.F. 1990b. Presynaptic mechanism for long-term potentiation in the hippocampus. *Nature* 346: 724–729.
- Bekkers, J.M. and Stevens, C.F. 1990c. Two different ways evolution makes neurons larger. In: *Understanding the Brain through the Hippocampus: The Hippocampal Region as a Model for Studying Brain Structure and Function*, edited by J. Storm-Mathisen, J. Zimmer, and O.P. Ottersen. Progress in Brain Research, vol. 83. Amsterdam: Elsevier, pp. 37–45.
- Bekkers, J.M., Richerson, G.B., and Stevens, C.F. 1990. Origin of variability in quantal size in cultured hippocampal neurons and hippocampal slices. *Proceedings of the National Academy of Sciences USA* 87: 5359–5362.
- Belardetti, F. and Siegelbaum, S.A. 1988. Up- and down-modulation of single K^+ channel function by distinct second messengers. *Trends in Neurosciences* 11: 232–238.

- Bernander, Ö, Douglas, R.J., Martin, K.A.C., and Koch, C. 1991. Synaptic background activity influences spatiotemporal integration in single pyramidal cells. *Proceedings of the National Academy of Sciences USA* 88: 11569–11573.
- Bernier, L., Castellucci, V.F., Kandel, E.R., and Schwartz, J.H. 1982. Facilitatory transmitter causes a selective and prolonged increase in adenosine 3',5'-monophosphate in sensory neurons mediating the gill and siphon withdrawal reflex in *Aplysia*. *Journal of Neuroscience* 2: 1682–1691.
- Bienenstock, E.L., Cooper, L.N., and Munro, P.W. 1982. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience* 2: 32–48.
- Bindman, L.J., Murphy, K.P., and Pockett, S. 1988. Postsynaptic control of the induction of long-term changes in efficacy of transmission at neocortical synapses in slices of rat brain. *Journal of Neurophysiology* 60: 1053–1065.
- Bittner, G.D. 1968. Differentiation of nerve terminals in the crayfish opener muscle and its functional significance. *Journal of General Physiology* 51: 731–758.
- Bittner, G.D. 1989. Synaptic plasticity at the crayfish opener neuromuscular preparation. *Journal of Neurobiology* 20: 386–408.
- Bizzi, E., Mussa-Ivaldi, F., and Giszter, S. 1991. Computations underlying the execution of movement: a biological perspective. *Science* 253: 287–291.
- Bland, B.H. 1986. The physiology and pharmacology of hippocampal formation theta rhythms. *Progress in Neurobiology* 26: 1–54.
- Bliss, T.V.P. and Collingridge, G.L. 1993. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361: 31–39.
- Bliss, T.V.P., Douglas, R.M., Errington, M.L., and Lynch, M.A. 1986. Correlation between long-term potentiation and release of endogenous amino acids from dentate gyrus of anaesthetized rats. *Journal of Physiology (London)* 377: 391–408.
- Bloedel, J.R., Bracha, V., Kelly, T.M., and Wu, J.-Z. 1991. Substrates for motor learning: does the cerebellum do it all? *Annals of the New York Academy of Sciences* 627: 305–318.
- Blumenfeld, H., Spira, M.E., Kandel, E.R., and Siegelbaum, S.A. 1990. Facilitatory and inhibitory transmitters modulate calcium influx during action potentials in *Aplysia* sensory neurons. *Neuron* 5: 487–499.
- Bok, S.T. 1959. *Histonomy of the Cerebral Cortex*. Amsterdam: Elsevier.
- Bonhoeffer, T., Staiger, V., and Aertsen, A. 1989. Synaptic plasticity in rat hippocampal slice cultures: local "Hebbian" conjunction of pre- and postsynaptic stimulation leads to distributed synaptic enhancement. *Proceedings of the National Academy of Sciences USA* 86: 8113–8117.
- Bower, J.M. 1990. Reverse engineering the nervous system: an anatomical, physiological, and computer-based approach. In: *An Introduction to Neural and Electronic Networks*, edited by S.F. Zornetzer, J.L. Davis, and C.L. Lau. San Diego, Calif.: Academic Press, pp. 3–22.
- Bradler, J.E. and Barrionuevo, G. 1989. Long-term potentiation in hippocampal CA3 neurons: tetanized inputs regulates heterosynaptic efficacy. *Synapse* 4: 132–142.
- Bradler, J.E. and Barrionuevo, G. 1990. Heterosynaptic correlates of long-term potentiation induction in hippocampal CA3 neurons. *Neuroscience* 35: 265–271.

- Braha, O., Dale, N., Hochner, B., Klein, M., Abrams, T.W., and Kandel, E.R. 1990. Second messengers involved in the two processes of presynaptic facilitation that contribute to sensitization and dishabituation in *Aplysia* sensory neurons. *Proceedings of the National Academy of Sciences USA* 87: 2040–2044.
- Breen, C. and Atwood, H.L. 1983. Octopamine---A neurohormone with presynaptic activity-dependent effects at crayfish neuromuscular junctions. *Nature* 303: 716–718.
- Bremermann, H. 1970. A method of unrestrained global optimization. *Mathematical Biosciences* 9: 1–15.
- Briggs, C.A. and McAfee, D.A. 1988. Long-term potentiation at nicotinic synapses in the rat superior cervical ganglion. *Journal of Physiology (London)* 404: 129–144.
- Brindley, G.S. 1967. The classification of modifiable synapses and their use in models of conditioning. *Proceedings of the Royal Society London B* 168: 361–376.
- Bröcher, S., Artola, A., and Singer, W. 1992. Intracellular injection of Ca^{2+} chelators blocks induction of long-term depression in rat visual cortex. *Proceedings of the National Academy of Sciences USA* 89: 123–127.
- Brodfehrer, P.D. and Friesen, W.O. 1986. Initiation of swimming activity by trigger neurons in the leech subesophageal ganglion. I. Output connections of Tr1 and Tr2. *Journal of Comparative Physiology A* 159: 489–502.
- Brotchie, P., Iansek, R., and Horne, M. 1991. A neural network model of neural activity in the monkey globus pallidus. *Neuroscience Letters* 131: 33–36.
- Brown, T.H., Chapman, P.F., Kairiss, E.W., and Keenan, C.L. 1988. Long-term synaptic potentiation. *Science* 242: 724–728.
- Brown, T.H., Ganong, A.H., Kairiss, E.W., Keenan, C.L., and Kelso, S.R. 1989. Long-term potentiation in two synaptic systems of the hippocampus brain slice. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 178–186.
- Brown, T.H., Kairiss, E.W., and Keenan, C.L. 1990. Hebbian synapses: biophysical mechanisms and algorithms. *Annual Review of Neuroscience* 13: 475–511.
- Brown, T.H., Zador, A.M., Mainen, Z.F., and Claiborne, B.J. 1991. Hebbian modifications in hippocampal neurons. In: *Long-Term Potentiation: A Debate of Current Issues*, edited by M. Baudry and J.L. Davis. Cambridge: MIT Press, pp. 357–389.
- Buchanan, J.T. 1992. Neural network simulations of coupled locomotor oscillators in the lamprey spinal cord. *Biological Cybernetics* 66: 367–374.
- Bullock, T. H. 1980. Reassessment of neural connectivity and its specification. In: *Information Processing in the Nervous System*, edited by H.M. Pinsky and W.D. Willis Jr. New York: Raven Press, pp. 199–220.
- Bulsara, A.R., Schieve, W.C., and Moss, F.E. 1992. Cooperative stochastic effects in a model of a single neuron. In: *Single Neuron Computation*, edited by T. McKenna, J. Davis, and S.F. Zornetzer. San Diego, CA: Academic Press, pp. 503–523.
- Buonomano, D.V. and Byrne, J.H. 1990. Long-term synaptic changes produced by a cellular analog of classical conditioning in *Aplysia*. *Science* 249: 420–423.
- Buonomano, D.V., Baxter, D.A., and Byrne, J.H. 1990. Small network of empirically derived adaptive elements simulate higher-order features of classical conditioning. *Neural Networks* 3: 507:523.
- Burke, W. 1966. Neuronal models for conditioned reflexes. *Nature* 210: 269–239.

- Burnod, Y. and Korn, H. 1989. Consequences of stochastic release of neurotransmitters for network computation in the central nervous system. *Proceedings of the National Academy of Sciences USA* 86: 352–356.
- Burrows, M. 1989. Processing of mechanosensory signals in local reflex pathways of the locust. *Journal of Experimental Biology* 146: 209–227.
- Burton, R.M., Jr. and Mpitsos, G.J. 1992. Event-dependent control of noise enhances learning in neural networks. *Neural Networks* 5: 627–637.
- Bush, P.C. and Douglas, R.J. 1991. Synchronization of bursting action potential discharge in a model network of neocortical neurons. *Neural Computation* 3: 19–30.
- Buxbaum, J.D. and Dudai, Y. 1989. Quantitative model for the kinetics of cAMP-dependent protein kinase (Type II) activity. *Journal of Biological Chemistry* 264: 9344–9351.
- Byrne, J.H. 1982. Analysis of synaptic depression contributing to habituation of gill-withdrawal reflex in *Aplysia californica*. *Journal of Neurophysiology* 48: 431–438.
- Byrne, J.H. 1987. Cellular analysis of associative learning. *Physiological Reviews* 67: 329–439.
- Byrne, J.H. and Crow, T. 1991. Examples of mechanistic analyses of learning and memory in invertebrates. In: *Learning and Memory: A Biological View*, edited by J.L. Martinez, Jr. and R.P. Kesner. San Diego, Calif.: Academic Press, pp. 329–358.
- Byrne, J.H. and Gingrich, K.J. 1989. Mathematical model of cellular and molecular processes contributing to associative and nonassociative learning in *Aplysia*. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 58–73.
- Byrne, J.H., Gingrich, K.J., and Baxter, D.A. 1989. Computational capabilities of single neurons: relationship to simple forms of associative and nonassociative learning in *Aplysia*. In: *Computational Models of Learning in Simple Systems*, edited by R.D. Hawkins and G.H. Bower. San Diego, Calif.: Academic Press, pp. 31–64.
- Byrne, J.H., Baxter, D.A., Buonomano, D.V., Cleary, L.J., Eskin, A., Goldsmith, J.R., McClendon, E., Nazif, F.A., Noel, F., and Scholz, K.P. 1991a. Neural and molecular bases of nonassociative and associative learning in *Aplysia*. *Annals of the New York Academy of Sciences* 627: 124–149.
- Byrne, J.H., Baxter, D.A., Buonomano, D.V., and Raymond, J.L. 1991b. Neuronal and network determinants of simple higher-order features of associative learning: experimental and modeling approaches. *Cold Spring Harbor Symposia on Quantitative Biology* 55: 165–186.
- Byrne, J.H., Cleary, L.J., and Baxter, D.A. 1991c. Aspects of the neural and molecular mechanisms of short-term sensitization in *Aplysia*: modulatory effects of serotonin and cAMP on duration of action potentials, excitability and membrane currents in tail sensory neurons. In: *The Biology of Memory*, edited by L.R. Squire and E. Lindenlaub. New York: F.K. Schattauer Verlag, pp. 7–28.
- Byrne, J.H., Zwartjes, R., Homayouni, R., Critz, S.D., and Eskin, A. 1993. Roles of second messenger pathways in neuronal plasticity and in learning and memory: insights gained from *Aplysia*. In: *Advances in Second Messenger and Phosphoprotein Research, Vol. 27*, edited by S. Shenolikar and A.C. Nairn. New York: Raven Press, pp. 47–108.
- Calvin, W.H. and Stevens, C.F. 1968. Synaptic noise and other sources of randomness in motorneuron interspike intervals. *Journal of Neurophysiology* 31: 574–587.
- Carrier, D. R. 1991. Conflict in the hypaxial musculo-skeletal system: documenting an evolutionary constraint. *American Zoologist* 31: 644–654.

- Castellucci, V.F. and Kandel, E.R. 1976. Presynaptic facilitation as a mechanism for behavioral sensitization in *Aplysia*. *Science* 194: 1176–1180.
- Castellucci, V.F., Pinsker, H., Kupfermann, I., and Kandel, E.R. 1970. Neuronal mechanisms of habituation and dishabituation of the gill-withdrawal reflex in *Aplysia*. *Science* 167: 1745–1748.
- Changeux, J.-P. and Danchin, A. 1976. Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. *Nature* 264: 705–711.
- Chapeau-Blondeau, F. and Chauvet, G. 1991. A neural network model of the cerebellar cortex performing dynamic associations. *Biological Cybernetics* 65: 267–280.
- Chapman, G. 1958. The hydrostatic skeleton in the invertebrates. *Biological Reviews of the Cambridge Philosophical Society* 33: 338–371.
- Chapman, G. 1975. Versatility of hydraulic systems. *Journal of Experimental Zoology* 194: 249–270.
- Chauvet, G. 1986. Habituation rules for a theory of the cerebellar cortex. *Biological Cybernetics* 55: 201–209.
- Chiel, H.J. and Beer, R.D. 1989. A lesion study of a heterogeneous artificial neural network for hexapod locomotion. *Proceedings of the International Joint Conference on Neural Networks* I: 407–414.
- Chiel, H.J., Kupfermann, I., and Weiss, K.R. 1988. An identified histaminergic neuron can modulate the outputs of buccal-cerebral interneurons in *Aplysia* via presynaptic inhibition. *Journal of Neuroscience* 8: 49–63.
- Chiel, H.J., Weiss, K.R., and Kupfermann, I. 1990. Multiple roles of a histaminergic afferent neuron in the feeding behavior of *Aplysia*. *Trends in Neurosciences* 13: 223–227.
- Chiel, H.J., Beer, R.D., Quinn, R.D., and Espenschied, K.S. 1992a. Robustness of a distributed neural network controller for locomotion in a hexapod robot. *IEEE Transactions on Robotics and Automation* 8: 293–303.
- Chiel, H.J., Crago, P., Mansour, J.M., and Hathi, K. 1992b. Biomechanics of a muscular hydrostat: a model of lapping by a reptilian tongue. *Biological Cybernetics* 67: 403–415.
- Choi, J.J., Oh, S., and Marks, R.J. 1991. Training layered perceptrons using low accuracy computation. *Proceedings 1991 IEEE International Conference on Neural Networks, Singapore* I: 554–559.
- Churchland, P.S. and Sejnowski, T.J. 1992. *The Computational Brain*. Cambridge, Mass.: MIT Press.
- Clark, G.A. and Kandel, E.R. 1984. Branch-specific heterosynaptic facilitation in *Aplysia* siphon sensory cells. *Proceedings of the National Academy of Sciences U.S.A.* 81: 2577–2581.
- Clark, G.A., Hawkins, R.D., and Frost, W.N. 1992. How neural are neural networks? A comparison of information processing and storage in artificial and real neural systems. *Journal of Statistical Planning and Inference* 33: 27–65.
- Cline, H.T. 1991. Activity-dependent plasticity in the visual system of frogs and fish. *Trends in Neurosciences* 14: 104–111.
- Clothiaux, E.E., Bear, M.F., and Cooper, L.N. 1991. Synaptic plasticity in visual cortex: comparison of theory with experiment. *Journal of Neurophysiology* 66: 1785–1804.
- Cohen, J.L., Weiss, K.R., and Kupfermann, I. 1978. Motor control of buccal muscles in *Aplysia*. *Journal of Neurophysiology* 41: 157–180.

- Collingridge, G.L. and Bliss, T.V.P. 1987. NMDA receptors—their role in long-term potentiation. *Trends in Neurosciences* 7: 288–293.
- Collingridge, G.L. and Singer, W. 1990. Excitatory amino-acid receptors and synaptic plasticity. *Trends in Pharmacological Sciences* 11: 290–296.
- Collingridge, G.L., Kehl, S.J., and McLennan, H. 1983. The antagonism of amino acid induced excitations of rat hippocampal CA1 neurones *in vitro*. *Journal of Physiology (London)* 334: 19–31.
- Collins, W.F. III, Honig, M.G., and Mendell, L.M. 1984. Heterogeneity of group Ia synapses on homonymous α -motoneurons as revealed by high-frequency stimulation of Ia afferent fibers. *Journal of Neurophysiology* 52: 980–991.
- Connor, J.A., Kretz, R., and Shapiro, E. 1986. Calcium levels measured in a presynaptic neurone of *Aplysia* under conditions that modulate transmitter release. *Journal of Physiology (London)* 375: 625–642.
- Cooper, L.N. 1986. Neuron learning to brain organization. *Cell Biophysics* 9: 103–144.
- Cooper, L.N., Lieberman, F., and Oja, E. 1979. A theory for the acquisition and loss of neuron specificity in visual cortex. *Biological Cybernetics* 33: 9–28.
- Cooper, L.N., Munro, P., and Scofield, C. 1985. Neuron selectivity: single neuron and neuron networks. In: *Synaptic Modification, Neuron Selectivity, and Nervous System Organization*, edited by W.B. Levy, J.A. Anderson, and S. Lehmkuhle. Hillsdale, N.J.: Lawrence Erlbaum, pp. 175–192.
- Cooper, L.N., Bear, M.F., Ebner, F.F., and Scofield, C. 1990. Neural networks: test tubes to theorems. In: *Brain Organization and Memory: Cells, Systems, and Circuits*, edited by J.L. McGaugh, N.M. Weinberger, and G. Lynch. New York: Oxford University Press, pp. 306–322.
- Crepel, F. and Jaillard, D. 1991. Pairing of pre- and postsynaptic activities in cerebellar Purkinje cells induces long-term changes in synaptic efficacy *in vitro*. *Journal of Physiology (London)* 432: 123–141.
- Crepel, F. and Krupa, M. 1988. Activation of protein kinase C induces a long-term depression of glutamate sensitivity of cerebellar Purkinje cells. *Brain Research* 458: 397–401.
- Crick, F. 1989. The recent excitement about neural networks. *Nature* 337: 129–132.
- Critz, S.D., Baxter, D.A., and Byrne, J.H. 1991. Modulatory effects of serotonin, FMRFamide, and myomodulin on the duration of action potentials, excitability and membrane currents in tail sensory neurons of *Aplysia*. *Journal of Neurophysiology* 66: 1912–1926.
- Cropper, E.C., Lloyd, P.E., Reed, W., Tenenbaum, R., Kupermann, I., and Weiss, K.R. 1987. Multiple neuropeptides in cholinergic motor neurons of *Aplysia*: evidence for modulation intrinsic to the motor circuit. *Proceedings of the National Academy of Sciences USA* 84: 3486–3490.
- Dale, N., Schacher, S., and Kandel, E.R. 1988. Long-term facilitation in *Aplysia* involves increase in transmitter release. *Science* 239: 282–285.
- Dan, Y. and Poo, M.-m. 1992. Hebbian depression of isolated neuromuscular synapses *in vitro*. *Science* 256: 1570–1573.
- Dayan, P. and Willshaw, D.J. 1991. Optimizing synaptic learning rules in linear associative memories. *Biological Cybernetics* 65: 253–265.
- Del Castillo, J. and Katz, B. 1954. Quantal components of the end-plate potential. *Journal of Physiology (London)* 124: 560–573.

- Desmond, J.E. and Moore, J.W. 1988. Adaptive timing in neural networks: the conditioned response. *Biological Cybernetics* 58: 405–415.
- Dolphin, A.C., Errington, M.L., and Bliss, T.V.P. 1982. Long-term potentiation of the perforant path *in vivo* is associated with increased glutamate release. *Nature* 297: 496–498.
- Donegan, N.H., Gluck, M.A., and Thompson, R.F. 1989. Integrating behavioral and biological models of classical conditioning. In: *Computational Models of Learning in Simple Neural Systems*, edited by R.D. Hawkins and G.H. Bower. San Diego, Calif.: Academic Press, pp. 109–156.
- Douglas, R.J. and Martin, K.A.C. 1991. Opening the grey box. *Trends in Neurosciences* 14: 286–293.
- Doya, K. and Yoshizawa, S. 1989. Adaptive neural oscillator using continuous-time back-propagation learning. *Neural Networks* 2: 375–385.
- Dudai, Y. 1987. The cAMP cascade in the nervous system: molecular sites of action and possible relevance to neuronal plasticity. *CRC Critical Reviews Biochemistry* 22: 221–281.
- Dudek, S.M. and Bear, M.F. 1992. Homosynaptic long-term depression in area CA1 of hippocampus and effects of *N*-methyl-D-aspartate receptor blockage. *Proceedings of the National Academy of Sciences USA* 89: 4363–4367.
- Dunwiddie, T. and Lynch, G. 1978. Long-term potentiation and depression of synaptic responses in the rat hippocampus: localization and frequency dependency. *Journal of Physiology (London)* 276: 353–367.
- Durbin, R. and Rumelhart, D.E. 1989. Product units: a computationally powerful and biologically plausible extension to backpropagation networks. *Neural Computation* 1: 133–142.
- Eccles, J.C. 1953. *The Neurophysiological Basis of Mind*. Oxford, U.K.: Clarendon, pp. 221–225.
- Edmonds, B., Klein, M., Dale, N., and Kandel, E.R. 1990. Contributions of two types of calcium channels to synaptic transmission and plasticity. *Science* 250: 1142–1147.
- Edwards, F.A., Konnerth, A., and Sakmann, B. 1990. Quantal analysis of inhibitory synaptic transmission in the dentate gyrus of rat hippocampal slices: a patch clamp study. *Journal of Physiology (London)* 430: 213–249.
- Eliot, L.S., Dudai, Y., Kandel, E.R., and Abrams, T.W. 1989. Ca^{2+} /calmodulin sensitivity may be common to all forms of neural adenylate cyclase. *Proceedings of the National Academy of Sciences USA* 86: 9564–9568.
- Erdos, P. and Niebur, E. 1990. The neural basis of the locomotion of nematodes. In: *Statistical Mechanics of Neural Networks*, edited by L. Garrido. Berlin: Springer-Verlag, pp. 253–267.
- Faggin, F. and Mead, C. 1990. VLSI implementation of neural networks. In: *An Introduction to Neural and Electronic Networks*, edited by S. F. Zornetzer, J.L. Davis, and C. Lau. San Diego, Calif.: Academic Press, pp. 275–292.
- Feinmark, S.J., Piomelli, D., Shapiro, E., and Schwartz, J.H. 1989. Biologically active metabolites of the 12-lipoxygenase pathway are formed by *Aplysia* nervous tissue. *Annals of the New York Academy of Sciences* 559: 121–130.
- Fetz, E.E. 1992. Are movement parameters recognizably coded in single neuron activity? *Behavioral and Brain Sciences* 15: 679–690.

- Fetz, E.E. and Shupe, L.E. 1990. Neural network models of the primate motor system. In: *Advanced Neural Computers*, edited by R. Eckmiller. Amsterdam: Elsevier North Holland, pp. 43–50.
- Fetz, E.E., Cheney, P.D., Mewes, K., and Palmer, S. 1989. Control of forelimb muscle activity by populations of corticomotoneuronal and rubromotoneuronal cells. *Progress in Brain Research* 80: 437–449.
- Fetz, E.E., Shupe, L.E., and Murthy, V. 1990. Neural networks controlling wrist movements. *Proceedings of the International Joint Conference on Neural Networks-90 II*: 675–679.
- Fiore, L. and Meunier, J.M. 1979. Synaptic connections and functional organization in *Aplysia* buccal ganglia. *Journal of Neurobiology* 10: 13–29.
- Flament, D., Fortier, P.A., and Fetz, E.E. 1992. Response patterns and postspike effects of peripheral afferents in dorsal root ganglia of behaving monkeys. *Journal of Neurophysiology* 67: 875–889.
- Fossier, P., Poulain, B., Baux, G., and Tauc, L. 1988. Both presynaptic nicotinic-like and muscarinic-like autoceptors regulate acetylcholine release at an identified neuro-neuronal synapse of *Aplysia*. *Pflügers Archiv* 411: 345–352.
- Frahm, H.D., Henz, S., and Baron, G. 1984. Comparison of brain structure volumes in insectivora and primates. V. Area striata (AS). *Journal für Hirnforschung* 25: 537–557.
- Frasconi, P., Gori, M., and Soda, G. 1992. Local feedback multilayered networks. *Neural Computation* 4: 120–130.
- Friedlander, M.J., Frégnac, Y., and Burke, J. 1993. Temporal covariance of post-synaptic membrane and synaptic input. Role in synaptic efficacy in visual cortex. *Progress in Brain Research* 95: 207–223.
- Friedrich, P. 1990. Protein structure: the primary substrate for memory. *Neuroscience* 35: 1–7.
- Friesen, W.O. 1985. Neuronal control of leech swimming movements: interactions between cell 60 and previously described oscillator neurons. *Journal of Comparative Physiology A* 156: 231–242.
- Friesen, W.O. 1989a. Neuronal control of leech swimming movements II. Motor neuron feedback to oscillator cells 115 and 28. *Journal of Comparative Physiology A* 166: 205–215.
- Friesen, W.O. 1989b. Neuronal control of leech swimming movements. In: *Neuronal and Cellular Oscillators*, edited by J.W. Jacklet. New York: Marcel Dekker, pp. 57–108.
- Fujii, S., Saito, K., Miyakawa, H., Ito, K., and Kato, H. 1991. Reversal of long-term potentiation (depotential) induced by tetanus stimulation of the input to CA1 neurons of guinea pig hippocampal slices. *Brain Research* 555: 112–122.
- Fuster, J.M. 1984. Behavioral electrophysiology of the prefrontal cortex. *Trends in Neurosciences* 7: 408–414.
- Fuster, J.M., Bauer, R.H., and Jervey, J.P. 1985. Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research* 330: 299–307.
- Gally, J.A., Montague, P.R., Reeke, G.N. Jr., and Edelman, G.M. 1990. The NO hypothesis: possible effects of a short-lived, rapidly diffusible signal in the development and function of the nervous system. *Proceedings of the National Academy of Sciences USA* 87: 3547–3551.
- Gamble, E. and Koch, C. 1987. The dynamics of free calcium in dendritic spines in response to repetitive synaptic input. *Science* 236: 1311–1315.

- Gardner, D. 1971. Bilateral symmetry and interneuronal organization in the buccal ganglia of *Aplysia*. *Science* 173: 550–553.
- Gardner, D. 1977a. Voltage-clamp analysis of a self-inhibitory synaptic potential in the buccal ganglia of *Aplysia*. *Journal of Physiology (London)* 264: 893–920.
- Gardner, D. 1977b. Interconnections of identified multi-action interneurons in the buccal ganglia of *Aplysia*. *Journal of Neurophysiology* 40: 349–361.
- Gardner, D. 1980a. Membrane-potential effects on an inhibitory post-synaptic conductance in *Aplysia* buccal ganglia. *Journal of Physiology (London)* 304: 165–180.
- Gardner, D. 1980b. Time integral of synaptic conductance. *Journal of Physiology (London)* 304: 181–191.
- Gardner, D. 1986. Variations in amplitude and time course of inhibitory postsynaptic currents. *Journal of Neurophysiology* 56: 1424–1438.
- Gardner, D. 1987. Synaptic diversity characterizes biological neural networks. *Proceedings of the IEEE First International Conference on Neural Networks IV*: 17–22.
- Gardner, D. 1989a. Sets of synaptic currents paired by common presynaptic or postsynaptic neurons. *Journal of Neurophysiology* 64: 845–853.
- Gardner, D. 1989b. Noise modulation of synaptic weights in a biological neural network. *Neural Networks* 2: 69–76.
- Gardner, D. 1990. Paired individual and mean postsynaptic currents recorded in four-cell networks of *Aplysia*. *Journal of Neurophysiology* 63: 1226–1240.
- Gardner, D. 1991. Presynaptic transmitter release is specified by postsynaptic neurons of *Aplysia* buccal ganglia. *Journal of Neurophysiology* 66: 2150–2154.
- Gardner, D. 1993a. Retrograde synaptic chemical transmitters could plausibly mediate neurobiological back-propagation. *Biophysical Journal* 64: A347.
- Gardner, D. 1993b. Transmitter release in *Aplysia*: applicability of quantal models and evidence for postsynaptic control. In: *Botulinum and Tetanus Neurotoxins: Neurotransmission and Biomedical Aspects*, edited by B.R. DasGupta. New York: Plenum Press, pp. 117–128.
- Gardner, D. and Kandel, E.R. 1972. Diphasic post synaptic potential: a chemical synapse capable of mediating conjoint excitation and inhibition. *Science* 176: 765–768.
- Gardner, D. and Kandel, E.R. 1977. Physiological and kinetic properties of cholinergic receptors activated by multi-action interneurons in the buccal ganglia of *Aplysia*. *Journal of Neurophysiology* 40: 333–348.
- Gardner, D. and Stevens, C.F. 1980. Rate-limiting step of inhibitory postsynaptic current decay in *Aplysia* buccal ganglia. *Journal of Physiology (London)* 304: 145–164.
- Gardner, D., Ruff, R.L., and White, R.L. 1984. Choline acts as agonist and blocker for *Aplysia* cholinergic synapses. *Journal of Neurophysiology* 51: 1–15.
- Garthwaite, J., Charles, S.L., and Chess-Williams, R. 1988. Endothelium-derived relaxing factor release on activation of NMDA receptors suggests role as intercellular messenger in the brain. *Nature* 336: 385–388.
- Geman, S., Bienenstock, E., and Doursat, R. 1992. Neural networks and the bias/variance dilemma. *Neural Computation* 4: 1–58.
- Gerschenfeld, H.M. and Paupardin-Tritsch, D. 1974. On the transmitter function of 5-hydroxytryptamine at excitatory and inhibitory monosynaptic junctions. *Journal of Physiology (London)* 243: 457–481.

- Gerstein, G.L. and Turner, M.R. 1990. Neural assemblies as building blocks of cortical computation. In: *Computational Neuroscience*, edited by E.L. Schwartz. Cambridge, Mass.: MIT Press, pp. 179–191.
- Getting, P.A. 1989. Emerging principles governing the operation of neural networks. *Annual Reviews of Neuroscience* 12: 185–204.
- Getting, P.A. and Dekin, M.S. 1985. *Tritonia* swimming: a model system for integration within rhythmic motor systems. In: *Model Neural Networks and Behavior*, edited by A.I. Selverston. New York: Plenum Press, pp. 3–20.
- Ghez, C. 1991. Muscles: Effectors of the motor systems. In: *Principles of Neural Science*, edited by E.R. Kandel, J.H. Schwartz and T.M. Jessell. 3rd edition. New York: Elsevier, pp. 548–563.
- Gilbert, C.D. and Wiesel, T.N. 1983. Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience* 3: 1116–1133.
- Gingrich, K.J. and Byrne, J.H. 1984. Simulation of nonassociative and associative neuronal modifications in *Aplysia*. *Society for Neuroscience Abstracts* 10: 270.
- Gingrich, K.J. and Byrne, J.H. 1985. Simulation of synaptic depression, posttetanic potentiation, and presynaptic facilitation of synaptic potentials from sensory neurons mediating the gill-withdrawal reflex in *Aplysia*. *Journal of Neurophysiology* 53: 652–669.
- Gingrich, K.J. and Byrne, J.H. 1987. Single-cell model for associative learning. *Journal of Neurophysiology* 57: 1705–1715.
- Gingrich, K.J., Baxter, D.A., and Byrne, J.H. 1988. Mathematical model of cellular mechanisms contributing to presynaptic facilitation. *Brain Research Bulletin* 21: 513–520.
- Glanzman, D.L., Kandel, E.R., and Schacher, S. 1989. Identified target motor neuron regulates neurite outgrowth and synapse formation of *Aplysia* sensory neurons in vitro. *Neuron* 3: 441–450.
- Glanzman, D.L., Kandel, E.R., and Schacher, S. 1990. Target-dependent structural changes accompanying long-term synaptic facilitation in *Aplysia* neurons. *Science* 249: 799–802.
- Gluck M.A. and Thompson, R.F. 1987. Modeling the neural substrates of associative learning and memory: a computational approach. *Psychology Review* 94: 176–191.
- Gnad, J.W. and Andersen, R.A. 1988. Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research* 70: 215–220.
- Goldman, R.S., Chavez-Noriega, L.E., and Stevens, C.F. 1990. Failure to reverse long-term potentiation by coupling sustained presynaptic activity and N-methyl-D-aspartate receptor blockade. *Proceedings of the National Academy of Sciences USA* 87: 7165–7169.
- Goldstein, R., Kistler, H.B. Jr., Steinbusch, H.W.M., and Schwartz, J.H. 1984. Distribution of serotonin-immunoreactivity in juvenile *Aplysia*. *Neuroscience* 11: 535–547.
- Graham, D. 1977. Simulation of a model for the coordination of leg movement in free walking insects. *Biological Cybernetics* 26: 187–198.
- Granzow, B., Friesen, W.O., and Kristan, W.B. Jr. 1985. Physiological and morphological analysis of synaptic transmission between leech motor neurons. *Journal of Neuroscience* 5: 2035–2050.
- Greenberg, S.M., Castellucci, V.F., Bayley, H., and Schwartz, J.H. 1987. A molecular model for long-term sensitization in *Aplysia*. *Nature* 329: 62–65.
- Greenough, W.T. and Anderson, B.J. 1991. Cerebellar synaptic plasticity: relation to learning versus neural activity. *Annals of the New York Academy of Sciences* 627: 231–247.

- Greuel, J.M., Luhmann, H.J., and Singer, W. 1988. Pharmacological induction of use-dependent receptive field modifications in the visual cortex. *Science* 242: 74–77.
- Grillner, S., Wallen, P., and Brodin, L. 1991. Neuronal network generating locomotor behavior in lamprey: circuitry, transmitters, membrane properties and simulation. *Annual Review of Neuroscience* 8: 263–305.
- Grover, L.M. and Teyler, T.J. 1990. Two components of long-term potentiation induced by different patterns of afferent activation. *Nature* 347: 447–479.
- Gustafsson, B. and Wigström, H. 1986. Hippocampal long-lasting potentiation produced by pairing single volley and brief conditioning tetani evoked in separate afferents. *Journal of Neuroscience* 6: 1575–1582.
- Gustafsson, B. and Wigström, H. 1988. Physiological mechanisms underlying long-term potentiation. *Trends in Neurosciences* 11: 156–162.
- Gustafsson, B., Wigström, H., Abraham, W.C., and Huang, Y.-Y. 1987. Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *Journal of Neuroscience* 7: 774–780.
- Hardesty, I. 1902. Observations on the medulla spinalis of the elephant with some comparative studies of the intumescencia cervicalis and the neurones of the columna anterior. *Journal of Comparative Neurology* 12: 125–182.
- Harris-Warrick, R. M. and Marder, E. 1991. Modulation of neural networks for behavior. *Annual Review of Neuroscience* 14: 39–57.
- Hashemzadeh, G.H., Colbert, C.M., and Levy, W.B. 1991. Refining the temporal definition of an association at the neuronal level using long-term potentiation and long-term depression in the dentate gyrus. *Neuroscience Letters* 126: 37–40.
- Hawkins, R.D. 1989a. A biologically based computational model for several simple forms of learning. In: *Computational Models of Learning in Simple Neural Systems*, edited by R.D. Hawkins and G.H. Bower. San Diego, Calif.: Academic Press, pp. 65–108.
- Hawkins, R.D. 1989b. A simple circuit model for high-order features of classical conditioning in *Aplysia*. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 74–93.
- Hawkins, R.D. and Kandel, E.R. 1984. Is there a cell biological alphabet for simple forms of learning? *Psychological Review* 91: 375–391.
- Hawkins, R.D., Abrams T.W., Carew, T.J., and Kandel, E.R. 1983. A cellular mechanism of classical conditioning in *Aplysia*: activity-dependent amplification of presynaptic facilitation. *Science* 219: 400–405.
- Hebb, D.O. 1949. *The Organization of Behavior*. New York: Wiley, p. 62.
- Hille, B. 1992. *Ionic Channels of Excitable Membranes*. 2nd edition. Sunderland, Mass.: Sinauer Associates.
- Hinton, G.E. 1989. Connectionist learning procedures. *Artificial Intelligence* 40: 185–234.
- Hirano, T. 1990a. Depression and potentiation of the synaptic transmission between a granule cell and a Purkinje cell in rat cerebellar culture. *Neuroscience Letters* 119: 141–144.
- Hirano, T. 1990b. Effects of postsynaptic depolarization in the induction of synaptic depression between a granule cell and a Purkinje cell in rat cerebellar culture. *Neuroscience Letters* 119: 145–147.
- Hirano, T. 1990c. Synaptic transmission between rat inferior olivary neurons and cerebellar Purkinje cells in culture. *Journal of Neurophysiology* 63: 181–189.

- Hirano, T. and Hagiwara, S. 1988. Synaptic transmission between rat cerebellar granule and Purkinje cell in dissociated cell culture: effects of excitatory-amino acid transmitter antagonists. *Proceedings of the National Academy of Sciences USA* 85: 934–938.
- Hirsch, J.C. and Crepel, F. 1990. Use-dependent changes in synaptic efficacy in rat prefrontal neurons *in vitro*. *Journal of Physiology (London)* 427: 31–49.
- Hirsch, J.C. and Crepel, F. 1992. Postsynaptic calcium is necessary for the induction of LTP and LTD of monosynaptic EPSPs in prefrontal neurons. An *in vitro* study in the rat. *Synapse* 10: 173–175.
- Hochner, B., Klein, M., Schacher, S., and Kandel, E.R. 1986. Additional components in the cellular mechanisms of presynaptic facilitation contribute to behavioral dishabituation in *Aplysia*. *Proceedings of the National Academy of Sciences USA* 83: 8794–8798.
- Holmes, W.R. and Levy, W.B. 1990. Insight into associative long-term potentiation from computational models of NMDA receptor-mediated calcium influx and intracellular calcium concentration changes. *Journal of Neurophysiology* 63: 1148–1168.
- Hopfield, J.J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA* 79: 2554–2558.
- Hopkins, W.F. and Johnston, D. 1984. Frequency-dependent noradrenergic modulation of long-term potentiation in the hippocampus. *Science* 226: 350–352.
- Hopkins, W.F. and Johnston, D. 1988. Noradrenergic enhancement of long-term potentiation at mossy fiber synapses in the hippocampus. *Journal of Neurophysiology* 59: 667–687.
- Houk, J.C., Singh, S.P., Fisher, C., and Barto, G. 1990. An adaptive sensorimotor network inspired by the anatomy and physiology of the cerebellum. In: *Neural Networks for Control*, edited by W.T. Miller, R.S. Sutton, and P.J. Werbos. Cambridge, Mass.: MIT Press, pp. 301–348.
- Huang, Y.-Y., Colino, A., Selig, D.K., and Malenka, R.C. 1992. The influence of prior synaptic activity on the induction of long-term potentiation. *Science* 255: 730–733.
- Intrator, N. and Cooper, L.N. 1992. Objective function formulation of the BCM theory of visual cortical plasticity: statistical connections, stability conditions. *Neural Networks* 5: 3–17.
- Ito, M. 1989a. Long-term depression: possible cellular mechanism for learning mediated by the cerebellum. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 178–186.
- Ito, M. 1989b. Long-term depression. *Annual Review of Neuroscience* 12: 85–102.
- Ito, M. 1991. The cellular basis of cerebellar plasticity. *Current Opinion in Neurobiology* 1: 616–620.
- Ito, M. 1992. Long-term depression in the cerebellum, hippocampus and neocortex. In: *Encyclopedia of Learning and Memory*, edited by L. Squire. New York: Macmillan, pp. 352–357.
- Jaffe, D. and Johnston, D. 1990. Induction of long-term potentiation at hippocampal mossy-fiber synapses follows a Hebbian rule. *Journal of Neurophysiology* 64: 948–960.
- Jahan-Parwar, B., Wilson, A.H., and Fredman, S.M. 1983. Role of proprioceptive reflexes in control of feeding muscles of *Aplysia*. *Journal of Neurophysiology* 49: 1469–1480.
- Jahr, C.E. and Stevens, C.F. 1990a. A quantitative description of NMDA receptor-channel kinetic behavior. *Journal of Neuroscience* 10: 1830–1837.

- Jahr, C.E. and Stevens, C.F. 1990b. Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *Journal of Neuroscience* 10: 3178–3182.
- Johnston, D., Hopkins, W.F., and Gray, R. 1989. The role of norepinephrine in long-term potentiation at mossy-fiber synapses in the hippocampus. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 307–328.
- Kairiss, E.W., Keenan, C.L., and Brown, T.H. 1988. A biologically plausible implementation of a Hebbian covariance algorithm. *Society for Neuroscience Abstracts* 14: 832.
- Kandel, E.R. 1976. *Cellular Basis of Behavior*. San Francisco: Freeman.
- Kandel, E.R. and O'Dell, T.J. 1992. Are adult learning mechanisms also used for development? *Science* 258: 243–245.
- Kandel, E.R. and Schwartz, J.H. 1982. Molecular biology of learning: modulation of transmitter release. *Science* 281: 433–442.
- Kandel, E.R., Klein, M., Hochner, B., Shuster, M., Siegelbaum, S.A., Hawkins, R.D., Glanzman, D.L., Castellucci, V.F., and Abrams, T.W. 1987. Synaptic modulation and learning. In: *Synaptic Function*, edited by G.M. Edelman, W.E. Gall, and W.M. Cowan. New York: John Wiley, pp. 471–518.
- Kano, M. and Kato, M. 1987. Quisqualate receptors are specifically involved in cerebellar synaptic plasticity. *Nature* 325: 276–279.
- Katz, P.S. and Harris-Warrick, R.M. 1990. Neuromodulation of the crab pyloric central pattern generator by serotonergic/cholinergic proprioceptive afferents. *Journal of Neuroscience* 10: 1495–1512.
- Keenan, C.L., Baxter, D.A., and Brown, T.H. 1987. Multiplicative interaction between long-term potentiation and short-term facilitation in crayfish neuromuscular synapses. *Society for Neuroscience Abstracts* 13: 1233.
- Kelso, S.R., Gangong, A.H., and Brown, T.H. 1986. Hebbian synapses in hippocampus. *Proceedings of the National Academy of Sciences USA* 83: 5326–5330.
- Kenyon, G.T., Fetz, E.E., and Puff, R.D. 1990. Effects of firing synchrony on signal propagation in layered networks. In: *Advances in Neural Information Processing Systems 2*, edited by D.S. Touretzky. San Mateo, Calif.: Morgan Kaufmann, pp. 141–148.
- Kiehn, O. 1991. Plateau potentials and active integration in the 'final common pathway' for motor behaviour. *Trends in Neurosciences* 14: 68–73.
- Kier, W. M. and Smith, K. K. 1985. Tongues, tentacles, and trunks: the biomechanics of movement in muscular-hydrostats. *Zoological Journal of the Linnean Society* 83: 307–324.
- Kimura, F., Tsumoto, T., Nishigori, A., and Yoshimura, Y. 1990. Long-term depression but not potentiation is induced in Ca^{2+} -chelated visual cortex neurons. *Neuroreport* 1: 65–68.
- Kirk, M.D. 1989. Premotor neurons in the feeding system of *Aplysia californica*. *Journal of Neurobiology* 20: 497–512.
- Kirkpatrick, S., Gelatt, C.D. Jr., and Vecchi, M.P. 1983. Optimization by simulated annealing. *Science* 220: 671–680.
- Kitajima, T. and Hara, K. 1990. A model of the mechanisms of long-term potentiation in the hippocampus. *Biological Cybernetics* 64: 33–39.
- Kitajima, T. and Hara, K. 1991. A model of the mechanism of cooperativity and associativity of long-term potentiation in the hippocampus: a fundamental mechanism of associative memory and learning. *Biological Cybernetics* 64: 365–371.

- Klein, M. and Kandel, E.R. 1980. Mechanism of calcium current modulation underlying presynaptic facilitation and behavioral sensitization in *Aplysia*. *Proceedings of the National Academy of Sciences USA* 77: 6912–6916.
- Klopf, A.H. 1988. A neuronal model of classical conditioning. *Psychobiology* 16: 85–125.
- Klopf, A.H. 1989. Classical conditioning phenomena predicted by a drive-reinforcement model of neuronal function. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 104–132.
- Koch, C., Zador, A., and Brown, T.H. 1992. Dendritic spines: convergence of theory and experiment. *Science* 256: 973–974.
- Kohonen, T., 1984. *Self-Organization and Associative Memory*. Heidelberg: Springer-Verlag.
- Kohonen, T. 1990. Notes on neural computing and associative memory. In: *Brain Organization and Memory: Cells, Systems, and Circuits*, edited by J.L. McGaugh, N.M. Weinberger, and G. Lynch. New York: Oxford University Press, pp. 323–337.
- Konnerth, A. Llano, I., and Armstrong, C.M. 1990. Synaptic currents in cerebellar Purkinje cells. *Proceedings of the National Academy of Sciences USA* 87: 2662–2665.
- Koyano, K., Kuba, K., and Minota, S. 1985. Long-term potentiation of transmitter release induced by repetitive presynaptic activities in bull-frog sympathetic ganglia. *Journal of Physiology (London)* 359: 219–233.
- Kreiner, T., Kirk, M.D., and Scheller, R.H. 1987. Cellular and synaptic morphology of a feeding motor circuit in *Aplysia californica*. *Journal of Comparative Neurology* 264: 311–325.
- Kristan, W.B., Jr. 1980. Generation of rhythmic motor patterns. In: *Information Processing in the Nervous System*, edited by H.M. Pinsky and W.D. Willis. New York: Raven Press, pp. 241–261.
- Kristan, W.B., Jr. 1982. Sensory and motor neurons responsible for the local bending response in leeches. *Journal of Experimental Biology* 96: 161–180.
- Kumamoto, E. and Kuba, K. 1983. Sustained rise in ACh sensitivity of a sympathetic ganglion induced by post-synaptic electrical activities. *Nature* 305: 145–146.
- Larkman, A., Stratford, K., and Jack, J. 1991. Quantal analysis of excitatory synaptic action and depression in hippocampal slices. *Nature* 350: 344–347.
- Larkman, A., Hannay, T., Stratford, K., and Jack, J. 1992. Presynaptic release probability influences the locus of long-term potentiation. *Nature* 360: 70–73.
- Larson, J. and Lynch, G. 1986. Patterned stimulation at the theta frequency is optimal for induction of long-term potentiation. *Brain Research* 368: 7–35.
- LeCun, Y. Denker, J. Solla, S., Howard, R.E., and Jackel, L.D. 1990. Optimal brain damage. In: *Advances in Neural Information Processing Systems 2*, edited by D.S. Touretzky. San Mateo, Calif.: Morgan Kaufmann, pp. 598–605.
- Lehky, S. and Sejnowski, T.J. 1988. Network model of shape-from-shading: neural function arises from both receptive and projective fields. *Nature* 333: 452–454.
- Lehky, S.R. and Sejnowski, T.J. 1990. Neural network model of visual cortex for determining surface curvature from images of shaded surfaces. *Proceedings of the Royal Society London B* 240: 251–278.
- Lehky, S.R., Sejnowski, T.J., and Desimone, R. 1992. Predicting responses of nonlinear neurons in monkey striate cortex to complex patterns. *Journal of Neuroscience* 12: 3568–3581.

- LeVay, S. 1988. The patchy intrinsic projections of visual cortex. *Progress in Brain Research* 75: 147–161.
- Levy, W.B. 1985. Associative changes at the synapse: LTP in the hippocampus. In: *Synaptic Modification, Neuron Selectivity, and Nervous System Organization*, edited by W.B. Levy, J.A. Anderson, and S. Lehmkuhle. Hillsdale, N.J.: Lawrence Erlbaum, pp. 5–33.
- Levy, W.B. 1989. A computational approach to hippocampal function. In: *Computational Models of Learning and Simple Neural Systems*, edited by R.D. Hawkins and G.H. Bower. San Diego, Calif.: Academic Press, pp. 243–306.
- Levy, W.B. and Colbert, C.M. 1991. Adaptive synaptogenesis can complement associative potentiation/depression. In: *Neural Network Models of Conditioning and Action*, edited by M.L. Commons, S. Grossberg, and J.E.R. Staddon. Hillsdale, N.J.: Lawrence Erlbaum, pp. 53–68.
- Levy, W.B. and Desmond, N.L. 1985. The rules of elemental synaptic plasticity. In: *Synaptic Modification, Neuron Selectivity, and Nervous System Organization*, edited by W.B. Levy, J.A. Anderson, and S. Lehmkuhle. Hillsdale, N.J.: Lawrence Erlbaum, pp. 105–121.
- Levy, W.B. and Steward, O. 1979. Synapses as associative memory elements in the hippocampal formation. *Brain Research* 175: 233–245.
- Levy, W.B. and Steward, O. 1983. Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience* 8: 791–797.
- Levy, W.B., Brassel, S.E., and Moore, S.D. 1983. Partial quantification of the associative synaptic learning rule of the dentate gyrus. *Neuroscience* 8: 799–808.
- Linden, D.J. and Connor, J.A. 1991. Participation of postsynaptic PKC in cerebellar long-term depression in culture. *Science* 254: 1656–1659.
- Linden, D.J., Dickinson, M.H., Smeyne, M., and Connor, J.A. 1991. A long-term depression of AMPA currents in cultured cerebellar Purkinje cells. *Neuron* 7: 81–89.
- Lisberger, S.G. and Sejnowski, T.J. 1992a. Computational analysis suggests a new hypothesis for motor learning in the vestibulo-ocular reflex. Technical Report INC-9201, Institute for Neural Computation, U.C. San Diego.
- Lisberger, S.G. and Sejnowski, T.J. 1992b. Motor learning in a recurrent network model based on the vestibulo-ocular reflex. *Nature* 360: 159–161.
- Lisman, J. 1989. A mechanism for the Hebbian and the anti-Hebbian processes underlying learning and memory. *Proceedings of the National Academy of Sciences USA* 86: 9574–9578.
- Livingstone, M.S. and Hubel, D.H. 1984. Specificity of intrinsic connections in primary visual cortex. *Journal of Neuroscience* 4: 2830–2835.
- Llinás, R.R. 1988. The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* 242: 1654–1664.
- Lloyd, P.E., Mahon, A.C., Kupfermann, I., Cohen, J.L., Scheller, R.H., and Weiss, K.R. 1985. Biochemical and immunocytological localization of molluscan small cardioactive peptides in the nervous system of *Aplysia californica*. *Journal of Neuroscience* 5: 1851–1861.
- Lo, Y.-J. and Poo, M.-m. 1991. Activity-dependent synaptic competition in vitro: heterosynaptic suppression of developing synapses. *Science* 254: 1019–1022.
- Lockery, S.R. 1989. *Distributed Processing of Sensory Information in the Leech*. Ph.D. Thesis, University of California, San Diego.

- Lockery, S.R. and Kristan, W.B. Jr. 1990a. Distributed processing of sensory information in the leech. I. Input-output relations of the local bending reflex. *Journal of Neuroscience* 10: 1811–1815.
- Lockery, S.R. and Kristan, W.B. Jr. 1990b. Distributed processing of sensory information in the leech. II. Identification of interneurons in the local bending reflex. *Journal of Neuroscience* 10: 1816–1829.
- Lockery, S.R. and Sejnowski, T.J. 1990. A dynamic neural network model of sensorimotor transformations in the leech. *Neural Computation* 2: 274–282.
- Lockery, S.R. and Sejnowski, T.J. 1991. Neural network models of chemotaxis in simple nervous systems. *Society for Neuroscience Abstracts* 17: 1490.
- Lockery, S.R. and Sejnowski, T.J. 1992. Distributed processing of sensory information in the leech III. A dynamical neural network model of the local bending reflex. *Journal of Neuroscience* 12: 3877–3895.
- Lockery, S.R., Wittenberg, G. Kristan, W.B. Jr., and Cottrell, G. 1989. Function of identified interneurons in the leech elucidated using neural networks trained by back-propagation. *Nature* 340: 648–471.
- Lockery, S.R., Fang, Y., and Sejnowski, T.J. 1990. A dynamical neural network model of sensorimotor transformations in the leech. *Neural Computation* 2: 274–282.
- Lopez, H.S., Burger, B., Dickstein, R., Desmond, N.L., and Levy, W.B. 1990. Associative synaptic potentiation and depression: quantification of dissociable modifications in the hippocampal dentate gyrus favors a particular class of synaptic modification equations. *Synapse* 5: 33–47.
- Lynch, G., Dunwiddie, T.V., and Gribkoff, V.K. 1977. Heterosynaptic depression: a postsynaptic correlate of long-term potentiation. *Nature* 266: 737–739.
- Lynch, G., Larson, J., Kelso, S., Barrionuevo, G., and Schottler, F. 1983. Intracellular injections of EGTA block the induction of hippocampal long-term potentiation. *Nature* 305: 719–721.
- Lytton, W.W. and Sejnowski, T.J. 1991. Simulations of cortical pyramidal neurons synchronized by inhibitory interneurons. *Journal of Neurophysiology* 66: 1059–1079.
- Madison, D.V., Malenka, R.C., and Nicoll, R.A. 1991. Mechanisms underlying long-term potentiation of synaptic transmission. *Annual Review of Neuroscience* 14: 379–397.
- Mahowald, M. and Douglas, R. 1991. A silicon neuron. *Nature* 354: 515–518.
- Malenka, R.C., Kauer, J.A., Zucker, R.S., and Nicoll, R.A. 1988. Postsynaptic calcium is sufficient for potentiation of hippocampal synaptic transmission. *Science* 242: 81–84.
- Malinow, R. 1991. Transmission between pairs of hippocampal slice neurons: quantal levels, oscillations, and LTP. *Science* 252: 722–724.
- Malinow, R. and Tsien, R.W. 1990. Presynaptic enhancement shown by whole-cell recordings of long-term potentiation in hippocampal slices. *Nature* 346: 177–180.
- Manabe, T., Renner, P., and Nicoll, R.A. 1992. Postsynaptic contribution to long-term potentiation revealed by the analysis of miniature synaptic currents. *Nature* 355: 50–55.
- Man-Son-Hing, H., Zoran, M.J., Ludowiak, K., and Haydon, P.G. 1989. A neuromodulator of synaptic transmission acts on secretory apparatus as well as on ion channels. *Nature* 341: 237–239.

- Marder, E. and Hooper, S.L. 1985. Neurotransmitter modulation of the stomatogastric ganglion of decapod crustaceans. In: *Model Neural Networks and Behavior*, edited by A.I. Selverston. New York: Plenum Press, pp. 319–337.
- Mayer, M.L., Westbrook, G.L., and Guthrie, P.B. 1984. Voltage-dependent block by Mg^{2+} of NMDA responses in spinal cord neurones. *Nature* 309: 261–263.
- Mayford, M., Barzilai, A., Keller, F., Schacher, S., and Kandel, E.R. 1992. Modulation of an NCAM-related adhesion molecule with long-term synaptic plasticity in *Aplysia*. *Science* 256: 638–644.
- Mazzoni, P., Andersen, R.A., and Jordan, M.I. 1991a. A more biologically plausible learning rule for neural networks. *Proceedings of the National Academy of Sciences USA* 88: 4433–4437.
- Mazzoni, P., Andersen, R.A., and Jordan, M.I. 1991b. A more biologically plausible learning rule than backpropagation applied to a network model of cortical area 7a. *Cerebral Cortex* 1: 293–307.
- McClellan, A.D. 1982a. Movements and motor patterns of the buccal mass of *Pleurobranchaea* during feeding, regurgitation, and rejection. *Journal of Experimental Biology* 98: 195–211.
- McClellan, A.D. 1982b. Re-examination of presumed feeding motor activity in the isolated nervous system of *Pleurobranchaea*. *Journal of Experimental Biology* 98: 212–228.
- McClelland, J.L. and Rumelhart, D.E. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*. Cambridge, Mass.: MIT Press.
- Miller, M.W., Lee, S.C., and Krasne, F.B. 1987. Cooperativity-dependent long-lasting potentiation in the crayfish lateral giant escape reaction circuit. *Journal of Neuroscience* 7: 1081–1092.
- Minsky, M.L. and Papert, S.A. 1969. *Perceptrons: An Introduction to Computational Geometry*. Cambridge, Mass.: MIT Press.
- Mishkin, M., Malamut, B., and Bachevalier, J. 1984. Memories and habits: two neural systems. In: *Neurobiology of Learning and Memory*, edited by G. Lynch, J.L. McGaugh, and N.M. Weinberger. New York: Guilford Press, pp. 65–77.
- Mjolsness, E., Sharp, D.H., and Reintz, J. 1991. A connectionist model of development. *Journal of Theoretical Biology* 152: 429–453.
- Moore, J.W. and Blazis, D.E.J. 1989. Simulation of a classically conditioned response: a cerebellar neural network implementation of the Sutton-Barto-Desmond model. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 187–207.
- Moore, J.W., Desmond, J.E., and Berthier, N.E. 1989. Adaptively timed conditioned responses and the cerebellum: a neural network approach. *Biological Cybernetics* 62: 17–28.
- Morris, R.G.M. 1990. Synaptic plasticity, neural architecture, and forms of memory. In: *Brain Organization and Memory: Cells, Systems, and Circuits*, edited by J.L. McGaugh, N.M. Weinberger, and G. Lynch. New York: Oxford University Press, pp. 52–77.
- Morton, D.W., Chiel, H.J., Cohen, L.B., and Wu, J.-y. 1991. Optical methods can be utilized to map the location and activity of putative motor neurons and interneurons during rhythmic activity in the buccal ganglion of *Aplysia*. *Brain Research* 564: 45–55.
- Mountcastle, V.B., Reitboeck, H.J., Poggio, G.F., and Steinmetz, M.A. 1992. Adaptation of the Reitboeck method of multiple microelectrode recording to the neocortex of the waking monkey. *Journal of Neuroscience Methods* 36: 77–84.

- Mozer, M.C. and Smolensky, P. 1989. Using relevance to reduce network size automatically. *Connection Science* 1: 3–16.
- Mulkey, R.M. and Malenka, R.C. 1992. Homosynaptic NMDA receptor-dependent long term depression in rat hippocampal slices. *Society for Neuroscience Abstracts* 18: 1498.
- Muller, K.J., Nicholls, J.G., and Stent, G.S., eds. 1981. *Neurobiology of the Leech*. Cold Spring Harbor: Cold Spring Harbor Laboratory.
- Munro, E., Shupe, L., and Fetz, E. 1991. Integration and differentiation in dynamic recurrent networks. *Proceedings 1991 International Conference on Neural Networks* II: 980.
- Nicholls, J.G. and Baylor, D.A. 1968. Specific modalities and receptive fields of sensory neurons in the CNS of the leech. *Journal of Neurophysiology* 31: 740–756.
- Nicoll, R.A., Kauer, J.A., and Malenka, R.C. 1988. The current excitement in long-term potentiation. *Neuron* 1: 97–103.
- Nowak, L., Bregestovski, P., Ascher, P., Herbert, A., and Prochiantz, A. 1984. Magnesium gates glutamate-activated channels in mouse central neurones. *Nature* 307: 462–465.
- Nunez, P.L. 1981. *Electric Fields of the Brain*. New York: Oxford University Press.
- Nusbaum, M.P., Weimann, J.M., Golowasch, J., and Marder, E. 1992. Presynaptic control of modulatory fibers by their neural network targets. *Journal of Neuroscience* 12: 2706–2714.
- Ocorr, K.A. and Byrne, J.H. 1985. Membrane responses and changes in cAMP levels in *Aplysia* sensory neurons produced by serotonin, tryptamine, FMRFamide, and small cardioactive peptide B (SCP_B). *Neuroscience Letters* 55: 113–118.
- Ocorr, K.A., Walters, E.T., and Byrne, J.H. 1985. Associative conditioning analog selectively increases cAMP levels of tail sensory neurons in *Aplysia*. *Proceedings of the National Academy of Sciences USA* 82: 2548–2552.
- O'Dell, T.J., Hawkins, R.D., Kandel, E.R., and Arancio, O. 1991. Tests of the roles of two diffusible substances in long-term potentiation: evidence for nitric oxide as a possible early retrograde messenger. *Proceedings of the National Academy of Sciences USA* 88: 11285–11289.
- Ono, J.K. 1989. Synaptic connections in the buccal ganglia of *Aplysia* mediated by an identified neuron containing a CCK/gastrin-like peptide co-localized with acetylcholine. *Brain Research* 493: 212–224.
- Ono, J.K. and McCaman, R.E. 1979. Measurement of endogenous transmitter levels after intracellular recording. *Brain Research* 165: 156–160.
- Ort, C.A., Kristan, W.B. Jr., and Stent, G.S. 1974. Neuronal control of swimming in the medicinal leech. II. Identification and connections of motor neurones. *Journal of Comparative Physiology A* 94: 121–154.
- Pahapill, P.A., Lnenicka, G.A., and Atwood, H.L. 1987. Long-term facilitation and low-frequency depression in a crayfish phasic motor axon. *Journal of Comparative Physiology A* 161: 367–375.
- Palmer, S.S. and Fetz, E.E. 1985. Discharge properties of primate forearm motor units during isometric muscle activity. *Journal of Neurophysiology* 54: 1178–1193.
- Pavrides, C., Greenstein, Y.J., Grudman, M., and Winson, J. 1988. Long-term potentiation in the dentate gyrus is induced preferentially on the positive phase of q-rhythm. *Brain Research* 439: 383–387.
- Pearlmutter, B.A. 1989. Learning state space trajectories in recurrent neural networks. *Neural Computation* 1: 263–269.

- Pearson, K.G. 1985. Are there central pattern generators for walking and flight in insects? In: *Feedback and Motor Control in Invertebrates and Vertebrates*, edited by W.J.P. Barnes and M.H. Gladden. London: Croom Helm, pp. 307–315.
- Pearson, K.G., Fourtner, C.R., and Wong, R.K. 1973. Nervous control of walking in the cockroach. In: *Control of Posture and Locomotion*, edited by R.B. Stein, K.G. Pearson, R.S. Smith, and J.B. Redford. New York: Plenum Press, pp. 495–514.
- Pearson, K.G., Reye, D.N., and Robertson, R.M. 1983. Phase-dependent influences of wing stretch receptors on flight rhythm in the locust. *Journal of Neurophysiology* 49: 1168–1181.
- Peter, N., Aranoff, B., Wu, F., and Schacher, S. 1992. Modulation of both an NCAM-like adhesion molecule on the surface of *Aplysia* neurons and neurite fasciculation by 5-HT and FMRFamide. *Society for Neuroscience Abstracts* 18: 1468.
- Pieroni, J.P. and Byrne, J.H. 1992. Differential effects of serotonin, FMRFamide, and small cardioactive peptide on multiple, distributed processes modulating sensorimotor synaptic transmission in *Aplysia*. *Journal of Neuroscience* 12: 2633–2647.
- Pineda, F.J. 1987. Generalization of backpropagation to recurrent neural networks. *Physical Review Letters* 18: 2229–2232.
- Pineda, F.J. 1989. Recurrent backpropagation and the dynamical approach to adaptive neural computation. *Neural Computation* 1: 161–172.
- Piomelli, D., Volterra, A., Dale, N., Siegelbaum, S.A., Kandel, E.R., Schwartz, J.H., and Belardetti, F. 1987. Lipxygenase metabolites of arachidonic acid as second messengers for presynaptic inhibition of *Aplysia* sensory cells. *Nature* 328: 38–43.
- Plummer, M.R. and Kirk, M.D. 1990. Premotor neurons B51 and B52 in the buccal ganglia of *Aplysia californica*: synaptic connections, effects on ongoing motor rhythms, and peptide modulation. *Journal of Neurophysiology* 63: 539–558.
- Pockett, S., Brookes, N.H., and Bindman, L.J. 1990. Long-term depression at synapses in slices of rat hippocampus can be induced by bursts of postsynaptic activity. *Experimental Brain Research* 80: 196–200.
- Prothero, J.W. and Sundsten, J.W. 1984. Folding of the cerebral cortex in mammals. A scaling model. *Brain Behavior and Evolution* 24: 152–167.
- Purves, D., Snider, W.D., and Voyvodic, J.T. 1988. Trophic regulation of nerve cell morphology and innervation in the autonomic nervous system. *Nature* 336: 123–128.
- Quinn, K.J., Schmajuk, N., Jain, A., Baker, J.F., and Peterson, B.W. 1992. Vestibuloocular reflex arc analysis using an experimentally constrained neural network. *Biological Cybernetics* 67: 113–22.
- Quinn, R.D. and Espenschied, K.S. 1993. Control of a hexapod robot using a biologically inspired neural network. In: *Biological Neural Networks in Invertebrate Neuroethology and Robotics*, edited by R.D. Beer, R.E. Ritzmann, and T. McKenna. San Diego, Calif.: Academic Press, pp. 365–381.
- Quintana, J., Fuster, J.M., and Yajeya, J. 1989. Effects of cooling parietal cortex on prefrontal units in delay tasks. *Brain Research* 503: 100–110.
- Raymond, J.L., Baxter, D.A., Buonomano, D.V., and Byrne, J.H. 1992. A learning rule based on empirically-derived activity-dependent neuromodulation supports operant conditioning in a small network. *Neural Networks* 5: 789–803.
- Redman, S. 1990. Quantal analysis of synaptic potentials in neurons of the central nervous system. *Physiological Reviews* 70: 165–198.

- Reilly, D.L. and Cooper, L.N. 1990. An overview of neural networks: early models to real world systems. In: *An Introduction to Neural and Electronic Networks*, edited by S.F. Zornetzer, J.L. Davis, and C. Lau. San Diego, Calif.: Academic Press, pp. 227–248.
- Reiter, H.O. and Stryker, M.P. 1988. Neural plasticity without postsynaptic action potentials: less-active inputs become dominant when kitten visual cortical cells are pharmacologically inhibited. *Proceedings of the National Academy of Sciences USA* 85: 3623–3627.
- Rescorla, R.A. and Wagner, A.R. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In: *Classical Conditioning II: Current Research and Theory*, edited by A.H. Black and W.F. Prokasy. New York: Appleton-Century-Crofts, pp. 64–99.
- Ringo, J.L. 1991. Neuronal interconnection as a function of brain size. *Brain Behavior and Evolution* 38: 1–6.
- Ritzmann, R.E., Pollack, A.J., Hudson, S.E., and Hyvonen, A. 1991. Convergence of multimodal sensory signals at thoracic interneurons of the escape system of the cockroach, *Periplaneta americana*. *Brain Research* 563: 175–183.
- Roberts, A. and Bush, B.M.H., eds. 1981. *Neurones without Impulses*. Cambridge, U.K.: Cambridge University Press.
- Rockel, A.J., Hirons, R.W., and Powell, T.P.S. 1980. The basic uniformity in structure of the neocortex. *Brain* 103: 221–244.
- Rockland, K.S. 1983. Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology* 216: 303–318.
- Rockland, K.S. and Lund, J.S. 1982. Widespread periodic intrinsic connections in the tree shrew visual cortex. *Science* 215: 1532–1534.
- Rosen, S.C., Weiss, K.L., Cohen, J.L., and Kupfermann, I. 1982. Interganglionic cerebral-buccal mechanoreceptors of *Aplysia*: receptive fields and synaptic connections to different classes of neurons involved in feeding behavior. *Journal of Neurophysiology* 48: 271–288.
- Rosen, S.C., Susswein, A.J., Cropper, E.C., Weiss, K.R., and Kupfermann, I. 1989. Selective modulation of spike duration by serotonin and the neuropeptides, FRMFamide, SCPB, buccalin and myomodulin in different classes of mechanoreceptor neurons in the cerebral ganglion of *Aplysia*. *Journal of Neuroscience* 9: 390–402.
- Rosen, S.C., Teyke, T., Miller, M.W., Weiss, K.R., and Kupfermann, I. 1991. Identification and characterization of cerebral-to-buccal interneurons implicated in the control of motor programs associated with feeding in *Aplysia*. *Journal of Neuroscience* 11: 3630–3655.
- Rosenblatt, F. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review* 65: 386–408.
- Rowat, P.F. and Selverston, A.I. 1991. Learning algorithms for oscillatory networks with gap junctions and membrane currents. *Network* 2: 17–42.
- Rubner, J. and Schulten, K. 1990. Development of feature detectors by self-organization. A network model. *Biological Cybernetics* 62: 193–199.
- Rumelhart, D.E. and McClelland, J.L., eds. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*. Cambridge, Mass.: MIT Press.
- Rumelhart, D.E., Hinton, G.E., and Williams, R.J. 1986a. Learning internal representations by back-propagating errors. *Nature* 323: 533–536.

- Rumelhart, D.E., Hinton, G.E., and Williams, R.J. 1986b. Learning internal representations by error propagation. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*. edited by D.E. Rumelhart and J.L. McClelland. Cambridge, Mass.: MIT Press, pp. 318–364.
- Rutishauser, U. and Landmesser, L. 1991. Polysialic acid on the surface of axons regulates patterns of normal and activity-dependent innervation. *Trends in Neurosciences* 14: 528–532.
- Sacktor, T.C. and Schwartz, J.H. 1990. Sensitizing stimuli cause translocation of protein kinase C in *Aplysia* sensory neurons. *Proceedings of the National Academy of Sciences USA* 87: 2036–2039.
- Sakurai, M. 1987. Synaptic modification of parallel-Purkinje cell transmission in *in vitro* guinea-pig slices. *Journal of Physiology (London)* 394: 463–480.
- Sakurai, M. 1990. Calcium is an intracellular mediator of the climbing fiber in induction of cerebellar long-term depression. *Proceedings of the National Academy of Sciences USA* 87: 3383–3385.
- Sawada, M., Wilkinson, J.M., MacAdoo, D.J., and Coggeshall, R.E. 1976. The identification of two inhibitory cells in each segmental ganglion of the leech and studies on the ionic mechanism of the inhibitory junctional potentials produced by these cells. *Journal of Neurobiology* 7: 435–445.
- Schacher, S., Montarolo, P.G., and Kandel, E.R. 1990. Selective short- and long-term effects of serotonin, small cardioactive peptide, and tetanic stimulation on sensorimotor synapses of *Aplysia* in culture. *Journal of Neuroscience* 10: 3286–3294.
- Schubert, D. 1991. The possible role of adhesion in synaptic modification. *Trends in Neurosciences* 14: 127–130.
- Schuman, E.M. and Madison, D.V. 1991. A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science* 254: 1503–1506.
- Schwartz, J.H. and Shkolnik, L.J. 1981. The giant serotonergic neuron of *Aplysia*: a multi-targeted nerve cell. *Journal of Neuroscience* 1: 606–619.
- Schwindt, P. C. 1992. Ionic currents governing input-output relations of Betz cells. In: *Single Neuron Computation*, edited by T. McKenna, J. Davis, and S.F. Zornetzer. San Diego, Calif.: Academic Press, pp. 235–258.
- Segev, I., Fleshman, J.W., and Burke, R.E. 1989. Compartmental models of complex neurons. In: *Methods in Neuronal Modeling*, edited by C. Koch and I. Segev. Cambridge, Mass.: MIT Press, pp. 63–96.
- Sejnowski, T.J. 1977. Storing covariance with nonlinearly interacting neurons. *Journal of Mathematical Biology* 4: 303–321.
- Sejnowski, T.J. and Stanton, P.K. 1990. Covariance storage in the hippocampus. In: *An Introduction to Neural and Electronic Networks*, edited by S.F. Zornetzer, J.L. Davis and C. Lau. San Diego, Calif.: Academic Press, pp. 365–377.
- Sejnowski, T.J. and Tesauero, G. 1989. The Hebbian rule for synaptic plasticity: algorithms and implementations. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 394–403.
- Sejnowski, T.J. and Tesauero, G. 1990. Building network learning algorithms from Hebbian synapses. In: *Brain Organization and Memory: Cells, Systems and Circuits*, edited by J.L. McGaugh, N.M. Weinberg, and G. Lynch. New York: Oxford University Press, pp. 338–355.

- Sejnowski, T.J., Koch, C., and Churchland, P.S. 1988. Computational neuroscience. *Science* 241: 1299–1306.
- Selverston, A.I. 1988. A consideration of invertebrate central pattern generators as computational data bases. *Neural Networks* 1: 109–117.
- Shapiro, E., Castellucci, V.F., and Kandel, E.R. 1980. Presynaptic membrane potential affects transmitter release in an identified neuron in *Aplysia* by modulating the Ca^{2+} and Mg^{2+} currents. *Proceedings of the National Academy of Sciences USA* 77: 629–633.
- Sharp, A.A., Abbott, L.F., and Marder, E. 1992. Artificial electrical synapses in oscillatory networks. *Journal of Neurophysiology* 67: 1691–1694.
- Shepherd, G.M. 1988. *Neurobiology*. New York: Oxford University Press.
- Shepherd, G.M. 1990a. *The Synaptic Organization of the Brain*. New York: Oxford University Press.
- Shepherd, G.M. 1990b. The significance of real neuron architectures for neural network simulations. In: *Computational Neuroscience*, edited by E.L. Schwartz. Cambridge, Mass.: MIT Press, pp. 82–96.
- Shimbel, A. 1950. Contributions to the mathematical biophysics of the central nervous system with special reference to learning. *Bulletin of Mathematical Biophysics* 12: 241–275.
- Siegelbaum, S.A. and Kandel, E.R. 1991. Learning-related synaptic plasticity: LTP and LTD. *Current Opinion in Neurobiology* 1: 113–120.
- Simonneau, M., Tauc, L., and Baux, G. 1980. Quantal release of acetylcholine examined by current fluctuation analysis at an identified neuro-neuronal synapse of *Aplysia*. *Proceedings of the National Academy of Sciences USA* 77: 1661–1665.
- Singer, W. 1983. Neural activity as an shaping factor in the self-organization of neuron assemblies. In: *Synergetics of the Brain*, edited by E. Basar, H. Flohr, H. Haken, and A.J. Mandell. New York: Springer-Verlag, pp. 89–101.
- Singer, W. 1987. Activity-dependent self-organization of synaptic connections as a substrate of learning. In: *The Neural and Molecular Bases of Learning*, edited by J.-P. Changeux and M. Konishi. New York: John Wiley, pp. 301–335.
- Singer, W. 1990. Search for coherence: a basic principle of cortical self-organization. *Concepts in Neuroscience* 1: 1–26.
- Singer, W., Gray, C., Engel, A., König, P., Artola, A., and Bröcher, S. 1990. Formation of cortical cell assemblies. *Cold Spring Harbor Symposia on Quantitative Biology* 55: 939–952.
- Skrede, K.K. and Malthe-Sørensen, D. 1981. Increased resting and evoked release of transmitter following repetitive electrical tetanization in hippocampus: a biochemical correlate to long-lasting synaptic potentiation. *Brain Research* 208: 436–441.
- Small, S.A., Kandel, E.R., and Hawkins, R.D. 1989. Activity-dependent enhancement of presynaptic inhibition in *Aplysia* sensory neurons. *Science* 243: 1603–1606.
- Small, S.A., O'Dell, T.J., Kandel, E.R., and Hawkins, R.D. 1992. Nitric oxide produces long-term enhancement of synaptic transmission in the CA1 region of hippocampus by an activity-dependent mechanism. *Society for Neuroscience Abstracts* 18: 639.
- Smith, A.W. and Zipser, D. 1989. Learning sequential structure with the real-time recurrent learning algorithm. *International Journal of Neural Systems* 1: 125–131.
- Smith, C.E. 1992. A heuristic approach to stochastic models of single neurons. In: *Single Neuron Computation*, edited by T. McKenna, J. Davis, and S.F. Zornetzer. San Diego, CA: Academic Press, pp. 561–588.

- Smith, D.O. 1975. Temporal characteristics of efferent neuron discharge during muscle contraction in the crayfish claw. *Journal of Comparative Physiology* 96: 273–283.
- Smith, K. K. 1984. The use of the tongue and hyoid apparatus during feeding in lizards (*Ctenosaura similis* and *Tupinambis nigropunctatus*). *Journal of Zoology* 202: 115–143.
- Sossin, W.S., Kirk, M.D., and Scheller, R.H. 1987. Peptidergic modulation of neuronal circuitry controlling feeding in *Aplysia*. *Journal of Neuroscience* 7: 671–681.
- Spain, W.J., Schwindt, P.C., and Crill, W.E. 1991. Post-inhibitory excitation and inhibition in layer V neurones from cat sensorimotor cortex. *Journal of Physiology (London)* 434: 609–626.
- Squire, L.R. 1987. *Memory and Brain*. New York: Oxford University Press.
- Stanton, P.K. and Sejnowski, T.J. 1989. Associative long-term depression in the hippocampus induced by hebbian covariance. *Nature* 339: 531–533.
- Stanton, P.K., Chattarji, S., and Sejnowski, T.J. 1989. Commissural synapses, but not mossy fiber synapses, in hippocampal field CA3 exhibit associative long-term potentiation and depression. *Brain Research* 495: 145–150.
- Stanton, P.K., Chattarji, S., and Sejnowski, T.J. 1991. 2-Amino-3-phosphonopropionic acid, an inhibitor of glutamate-stimulated phosphoinositide turnover, blocks induction of homosynaptic long-term depression, but not potentiation, in rat hippocampus. *Neuroscience Letters* 127: 61–66.
- Staubli, U. and Lynch, G. 1987. Stable hippocampal long-term potentiation elicited by “theta” pattern stimulation. *Brain Research* 435: 227–234.
- Staubli, U. and Lynch, G. 1990. Stable depression of potentiated synaptic responses in the hippocampus with 1–5 Hz stimulation. *Brain Research* 513: 113–118.
- Stein, P.S.G. 1983. The vertebrate scratch reflex. *Symposia of the Society for Experimental Biology* 37: 383–403.
- Stent, G.S. 1973. A physiological mechanism for Hebb’s postulate of learning. *Proceedings of the National Academy of Sciences USA* 70: 997–1001.
- Stevens, C.F. 1989. How cortical interconnectedness varies with network size. *Neural Computation* 1: 473–479.
- Stuart, A.E. 1970. Physiological and morphological properties of motor neurones in the central nervous system of the leech. *Journal of Physiology (London)* 209: 627–646.
- Sugita, S., Goldsmith, J.R., Baxter, D.A., and Byrne, J.H. 1992. Involvement of protein kinase C in serotonin-induced spike broadening and synaptic facilitation in sensorimotor connections of *Aplysia*. *Journal of Neurophysiology* 68: 643–651.
- Susswein, A.J. and Byrne, J.H. 1988. Identification and characterization of neurons initiating patterned neural activity in the buccal ganglia of *Aplysia*. *Journal of Neuroscience* 8: 2049–2061.
- Sutton, R.S. and Barto, A.G. 1981. Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review* 88: 135–170.
- Tank, D.W. 1989. What details of neural circuits matter? *Seminars in the Neurosciences* 1: 67–79.
- Taussig, R., Sweet-Cordero, A., and Scheller, R.H. 1989. Modulation of ionic currents in *Aplysia* motor neuron B15 by serotonin, neuropeptides, and second messengers. *Journal of Neuroscience* 9: 3218–3229.

- Tesauro, G. 1986. Simple neural models of classical conditioning. *Biological Cybernetics* 55: 187–200.
- Teyke, T., Weiss, K. R., and Kupfermann, I. 1990. Appetitive feeding behavior of *Aplysia*: behavioral and neural analysis of directed head turning. *Journal of Neuroscience* 10: 3922–3934.
- Teyler, T.J. and DiScenna, P. 1987. Long-term potentiation. *Annual Review of Neuroscience* 10: 131–161.
- Teyler, T., Aroniadou, V., Berry, R.L., Borroni, A., DiScenna, P., Grover, L., and Lambert, N. 1990. LTP in neocortex. *Seminars in the Neurosciences* 2: 365–379.
- Thompson, R.F. 1986. The neurobiology of learning and memory. *Science* 233: 941–947.
- Thompson, R.F. 1989. Neural circuit for classical conditioning of the eyelid closure response. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches*, edited by J.H. Byrne and W.O. Berry. San Diego, Calif.: Academic Press, pp. 160–177.
- Thompson, W.J. and Stent, G.S. 1976. Neuronal control of heartbeat in the medicinal leech. *Journal of Comparative Physiology A* 111: 309–333.
- Traub, R.D. and Wong, P.K.S. 1982. Cellular mechanism of synchronization in epilepsy. *Science* 216: 745–747.
- Traub, R.D., Miles, R., and Wong, P.K.S. 1989. Model of the origin of rhythmic population oscillations in the hippocampal slice. *Science* 243: 1319–1324.
- Tsumoto, T. 1990. Long-term potentiation and depression in the cerebral neocortex. *Japanese Journal of Physiology* 40: 573–593.
- Tsung, F.-S., Cottrell, G.W., and Selverston, A.I. 1990. Experiments on learning stable network oscillations. *Proceedings of the International Joint Conference on Neural Networks-90* I: 169–174.
- Vertosick, F.T. and Kelly, R.H. 1991. The immune system as a neural network: a multi-epitope approach. *Journal of Theoretical Biology* 150: 225–237.
- Wadepuhl, M. and Beyn, W.-J. 1989. Computer simulation of the hydrostatic skeleton. The physical equivalent, mathematics and application to worm-like forms. *Journal of Theoretical Biology* 136: 379–402.
- Wagner, A.R. and Donegan, N.H. 1989. Some relationships between a computational model (SOP) and a neural circuit for Pavlovian (rabbit eyeblink) conditioning. In: *Computational Models of Learning in Simple Neural Systems*, edited by R.D. Hawkins and G.H. Bower. San Diego, Calif.: Academic Press, pp. 157–204.
- Wainwright, S.A., Biggs, W.D., Currey, J.D., and Gosline, J.M. 1976. *Mechanical Design in Organisms*. New York: John Wiley.
- Walters, E.T. and Byrne, J.H. 1983. Associative conditioning of single sensory neurons suggests a cellular mechanism for learning. *Science* 219: 405–408.
- Walters, E.T. and Byrne, J.H. 1985. Long-term enhancement produced by activity-dependent modulation of *Aplysia* sensory neurons. *Journal of Neuroscience* 5: 662–672.
- Watrous, R.L. and Shastri, L. 1986. Learning phonetic features using connectionist networks: an experiment in speech recognition. *Technical Report MS-CIS-86-78*, Linc Lab 44, University of Pennsylvania, Philadelphia.
- Weimann, J.M., Meyrand, P., and Marder, E. 1991. Neurons that form multiple pattern generators: identification and multiple activity patterns of gastric/pyloric neurons in the crab stomatogastric system. *Journal of Neurophysiology* 65: 111–122.

- Weiss, K.R., Brezina, V., Cropper, E.C., Hooper, S.L., Miller, M.W., Probst, W.C., Vilim, F.S., and Kupermann, I. 1992. Peptidergic co-transmission in *Aplysia*: functional implications for rhythmic behaviors. *Experientia* 48: 456–463.
- Whim, M.D. and Lloyd, P.E. 1989. Frequency-dependent release of peptide cotransmitters from identified cholinergic motor neurons in *Aplysia*. *Proceedings of the National Academy of Sciences USA* 86: 9034–9038.
- Whim, M.D. and Lloyd, P.E. 1992. Modulation of peptide release from single identified *Aplysia* neurons in culture. *Journal of Neuroscience* 12: 3545–3553.
- White, G., Levy, W.B., and Steward, O. 1988. Evidence that associative interactions between synapses during the induction of long-term potentiation occur within local dendritic domains. *Proceedings of the National Academy of Sciences USA* 85: 2368–2372.
- White, G., Levy, W.B., and Steward, O. 1990. Spatial overlap between populations of synapses determines the extent of their associative interaction during the induction of long-term potentiation and depression. *Journal of Neurophysiology* 64: 1186–1189.
- Wickens, J. 1988. Electrically coupled but chemically isolated synapses: dendritic spines and calcium in a rule for synaptic modification. *Progress in Neurobiology* 31: 507–528.
- Wigström, H., Gustafsson, B., Huang, Y.-Y., and Abraham, W.C. 1986. Hippocampal long-term potentiation is induced by pairing single afferent volleys with intracellularly injected depolarizing pulses. *Acta Physiologica Scandinavica* 126: 317–319.
- Williams, J.H., Errington, M.L., Lynch, M.A., and Bliss, T.V.P. 1989. Arachidonic acid induces a long-term activity-dependent enhancement of synaptic transmission in the hippocampus. *Nature* 341: 739–742.
- Williams, R.J. and Zipser, D. 1989a. A learning algorithm for continually running fully recurrent neural networks. *Neural Computation* 1: 270–280.
- Williams, R.J. and Zipser, D. 1989b. Experimental analysis of the real-time recurrent learning algorithm. *Connection Science* 1: 87–111.
- Williams, R.J. and Zipser, D. 1990. Gradient-based learning algorithms for recurrent connectionist networks. Technical report NU-CCS-90-9, College of Computer Science, Northeastern University, Boston, Mass.
- Wilson, D.M. 1966. Insect walking. *Annual Review of Entomology* 11: 103–122.
- Wilson, J.A. 1979. *Principles of Animal Physiology*. New York: Macmillan.
- Wilson, J.F. 1984. Robotic mechanics and animal morphology. In: *Robotics and Artificial Intelligence*, edited by M. Brady, L. A. Gerhardt, and H. F. Davidson. Berlin: Springer-Verlag, pp. 419–443.
- Wilson, J.F., Mahajan, U., Wainwright, S.A., and Croner, L.J. 1991. A continuum model of elephant trunks. *Journal of Biomechanical Engineering* 113: 79–84.
- Wittenberg, G. and Kristan, W.B., Jr. 1992. Analysis and modeling of the multisegmental coordination of shortening behavior in the medicinal leech: II. Role of identified interneurons. *Journal of Neurophysiology* 68: 1693–1707.
- Woodson, P.B.J., Schlapfer, W.T., Tremblay, J.P., and Barondes, S.H. 1975. Cholinergic agents affect two receptors that modulate transmitter release at a central synapse in *Aplysia californica*. *Brain Research* 88: 455–474.
- Xie, X., Berger, T.W., and Barrionuevo, G. 1992. Isolated NMDA receptor-mediated synaptic responses express both LTP and LTD. *Journal of Neurophysiology* 67: 1009–1013.

- Yamada, W.M., Koch, C., and Adams, P.R. 1989. Multiple channels and calcium dynamics. In: *Methods in Neuronal Modeling*, edited by C. Koch and I. Segev. Cambridge, Mass.: MIT Press, pp. 97–133.
- Yang, X.-D. and Faber, D.S. 1991. Initial synaptic efficacy influences induction and expression of long-term changes in transmission. *Proceedings of the National Academy of Sciences USA* 88: 4299–4303.
- Yang, X.-D., Korn, H., and Faber, D.S. 1990. Long-term potentiation of electronic coupling at mixed synapses. *Nature* 348: 542–545.
- Yeo, C.H. 1991. Cerebellum and classical conditioning of motor responses. *Annals of the New York Academy of Sciences* 627: 292–303.
- Yovell, R. and Abrams, T.W. 1992. Temporal asymmetry in activation of *Aplysia* adenylyl cyclase by calcium and transmitter may explain temporal requirements of conditioning. *Proceedings of the National Academy of Sciences USA* 89: 6526–6530.
- Zador, A., Koch, C., and Brown, T.H. 1990. Biophysical model of a Hebbian synapse. *Proceedings of the National Academy of Sciences USA* 87: 6718–6722.
- Zalutsky, R.A. and Nicoll, R.A. 1990. Comparison of two forms of long-term potentiation in single hippocampal neurons. *Science* 248: 1619–1624.
- Zečević, D., Wu, J.-Y., Cohen, L.B., London, J.A., Höpp, H.-P., and Falk, C.X. 1989. Hundreds of neurons in the *Aplysia* abdominal ganglion are active during the gill-withdrawal reflex. *Journal of Neuroscience* 9: 3681–3689.
- Zill, S.N. and Moran, D.T. 1981. The exoskeleton and insect proprioception: III. Activity of tibial campaniform sensilla during walking in the american cockroach, *Periplaneta americana*. *Journal of Experimental Biology* 94: 57–75.
- Zipser, D. 1989. A subgrouping strategy that reduces complexity and speeds up learning in recurrent networks. *Neural Computation* 1: 551–557.
- Zipser, D. 1990. Short-term active memory: a recurrent network model of the neural mechanism. Report 9003, Department of Cognitive Science, University of California San Diego, San Diego, Calif.
- Zipser, D. 1991. Recurrent network model of the neural mechanism of short-term active memory. *Neural Computation* 3: 179–193.
- Zipser, D. 1992. Identification models of the nervous system. *Neuroscience* 47: 853–862.
- Zipser, D. and Andersen, R.A. 1988a. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* 331: 679–684.
- Zipser, D. and Andersen, R.A. 1988b. The role of the teacher in learning-based models of parietal area 7a. *Brain Research Bulletin* 21: 505–512.
- Zipser, D. and Rumelhart, D.E. 1990. The neurobiological significance of the new learning models. In: *Computational Neuroscience*, edited by E.L. Schwartz. Cambridge, Mass.: MIT Press, pp. 192–200.
- Zucker, R.S. 1989. Short-term synaptic plasticity. *Annual Review of Neuroscience* 12: 13–31.
- Zucker, R.S. and Bruner, J. 1977. Long-lasting depression and the depletion hypothesis at crayfish neuromuscular junctions. *Journal of Comparative Physiology* 121: 223–240.

Index

- Acetylcholine
 in *Aplysia*, 60
 in buccal ganglia of *Aplysia*, 30, 36, 60
 and retrosynaptic transmission, 60
- Activation function. *See* Neuronal input/output function
- Active zones, 61
- Activity signal, 68. *See also* Activity trace
- Activity trace, 86, 89, 99, 102
 Hebbian, 89
- Activity-dependent neuromodulation, 79
 in *Aplysia*, 99–102
 definition of, 98
 examples of, 98–102
 formalism describing, 99
 and LTP, 65
 mechanisms of, 99
 and nitric oxide, 69
 requirements for, 98
 and transmitter mobilization, 102
- Adenylyl cyclase. *See also* cAMP
 in activity-dependent
 neuromodulation, 101
- Algorithms, brain vs. network, 13
- Anti-Hebbian rules, 92–94
- Aplysia californica*
 buccal ganglia of. *See* Buccal ganglia of *Aplysia*
 defensive withdrawal reflexes of, 84
 feeding in, 30–31, 144
 histaminergic neurons of, 148
 sensorimotor synapses of, 76, 79, 99
- Arachidonic acid
 in *Aplysia*, 60, 84
 in hippocampus, 60, 63
- Attractors, 71
- Autapses, 36. *See also* Self-inhibition
- Backpropagation
 activity-coding presynaptic receptors
 for, 66–67
 as algorithm, 42, 111–112, 115
 credit assignment using, 2
 and discrete spiking, 188
 dual presynaptic receptors for, 65
 and error minimization, 27, 38–42, 111–112, 115, 167
 formalism describing, 38–40
 learning by, 17, 111–112
 nervous systems models using, 112–116
 neurobiological analogs to, 40, 41
 dual-pathway implementation for, 42
 neurobiological plausibility
 of, 6, 68, 115
 and oscillating networks, 170–172
 recurrent. *See also* Recurrent networks
 definition of, 112
 generating integral and differential, 184–187
 in local bending reflex model, 125, 129
 nervous system models using, 114–115
 and temporal flow model, 168
 retrosynaptic requirement for, 29, 65
 retrosynaptic transmission in, 40
 static, definition of, 112
 teacherless algorithms using, 41
 weight-coding presynaptic receptors
 for, 65–67
- Behaviorally correct patterns, 134
- Bilateral symmetry, 29
- Biological synaptic strengths
 aggregate, 36
 in buccal ganglia of *Aplysia*, 34–36
 dependence on geometry of, 140
 dynamic, 6, 17–19, 35
 and instantiations of networks, 33
 mechanism distinct from specification
 of, 49

- Biological synaptic strengths (*cont.*)
 - postsynaptic specification of, 46–48, 58
 - saturation of, 18
 - static, 17, 34
 - and synaptic weights, 7, 33
- Birds, 141
- Bizzi, E., 147
- Boltzmann machine, 41
- Bottom-up analysis, 7, 165, 170
- Brain
 - distributed processing in, 107
 - as parallel processor, 5
 - size of, 13–17
- Buccal ganglia of *Aplysia*
 - advantages of, 21
 - distributed processing in, 32
 - feeding control by, 30
 - firing patterns of, 30
 - four-cell network of, 43
 - identified cells of
 - analyses using, 22
 - B4 and B5, actions of, 30
 - S cell cluster of, 30
 - independent synaptic strengths in, 49
 - inputs to, 30
 - and network models, 32–33
 - neurotransmitters of, 37
 - overview of, 29
 - postsynaptic specification of presynaptic release, 50–64
 - postsynaptic specification of synaptic strength, 43–50
 - presynaptic mechanism for synaptic strength, 54–58
 - retrosynaptic transmission in, 58–70
 - self-inhibition in, 30
 - synaptic complexity of, 36
 - synaptic strengths in, 34–36
- Buccal musculature of *Aplysia*, 31, 148
- Buccalin
 - in buccal ganglia of *Aplysia*, 37
 - as heterosynaptic modulator, 61
- Bullfrog, 81
- Bullock, T.H., 140
- Ca²⁺
 - in activity-dependent neuromodulation, 99–102
 - as LTD/LTP switch, 92
- Ca²⁺ channels, 61
- Caenorhabditis elegans*, 143
- cAMP
 - in activity-dependent neuromodulation, 101
 - differing requirements for, 67
- cAMP-dependent protein kinase. *See* Protein kinase A
- Cats, 82, 141
 - brain size of, 13
 - muscle of, 145
- Cell-adhesion molecules, 61
- Centipedes, 141
- Central pattern generators, 114, 149, 169
 - and locomotion, 141, 161
- Cerebellum
 - flocculus of, 187
 - LTD in, 92
- Cerebral cortex. *See* Cortex
- Channels, membrane, 139
- Cholecystokinin
 - in buccal ganglia of *Aplysia*, 37
- Choline, 60
 - in buccal ganglia of *Aplysia*, 36
- Class one networks, 17
- Class two networks, 17–19, 33
- Classical conditioning, 104
- Cockroaches
 - escape response of, 137
 - kinematic model of, 155
 - locomotion of, 150
- Computational neuroethology, 162
- Conditional delay tasks, 169
- Conditioning, 104
- Conjoint synapses
 - and weight sign change, 33
- Connectionist networks, 22. *See also* Networks
 - analogies to neurobiology of, 23
 - and compartmental models, 110
 - and Hebbian synapses, 69
 - and LTP, 69
 - requirement for retrosynaptic transmission, 23
- Contractile biopolymers, 163. *See also* Deformable structures
- Cortex
 - cerebellar. *See* Cerebellum
 - entorhinal, 81, 86
 - intrinsic connections of, 16
 - laminae of. *See* Cortical laminae
 - models of, 108
 - modularity of, 14–17
 - neuronal density of, 14–17
 - parietal. *See* Parietal cortex
 - primary visual area of, 14–17. *See also* Visual cortex
 - scaling of, 13–17
 - synaptic density of, 14–15

- thalamic input to, 139
- waves of, 169
- Cortical laminae, thickness of, 16
- Cotransmission
 - in buccal ganglia of *Aplysia*, 37
- Covariance rule, 94–95
- Crayfish, 75–79, 79
 - activity-dependent neuromodulation in, 99
 - LTD in, 79
 - LTP in, 75–79
 - opener-neuromuscular synapse of, 75–79, 79
- Credit assignment, 2
- Crick, F.H.C., 6
- Decay time constant, synaptic, 33
- Deformable structures, 145
- Dekin, M.S., 140
- Depression, nonassociative, 81. *See also* Long-term depression
- Dihydropyridines, 61
- Diphasic synaptic potential, 30, 36
 - and weight sign change, 33
- Distributed process-
 - ing, 3, 32, 49, 107, 108
 - existence proof of, 131
- Dopamine, in buccal ganglia of *Aplysia*, 37
- Douglas, R.J., 7
- Dynamic recurrent networks, 7
 - applications for, 169–187
 - categories of, 169
 - definition of, 166–167
 - gated, 170–172
 - for neural integration and differentiation, 184–187
 - neuromorphic, 188–189
 - oscillating, 169–172
 - reduced versions of, 171, 184
 - for short-term memory, 180–184
 - for step-target tracking, 173–180
 - for vestibulo-ocular reflex, 184–187
- Ecological niches, 141
- Eels, 141
- Elastic materials, 145
- Electrotonic synapses, 36
 - models of, 126, 129
- Elephants, 144
- ELH (egg-laying hormone)
 - in buccal ganglia of *Aplysia*, 37
- EPSPs
 - in buccal ganglia of *Aplysia*, 30
 - in leech, 122
- Error minimization, 27–28, 38–40, 107, 167. *See also* Backpropagation
 - by Boltzmann machine, 41
 - by numerical differentiation, 114
 - by polynomial fits, 114
- Error signal, 39–40
 - backpropagation of, 40
 - weight change by, 40
- Evoked-potential recording, 4
- Evolution, 15, 137, 141, 163
- Exclusive-or, 2
- External teacher, requirement for, 28
- Feeding
 - in *Aplysia*, 30–31, 144
 - in humans, 142
- Field potentials, 139
- Field-potential recordings, 4
- Final common pathway, 148
- Finite state grammars, 169
- Fish, 141
- FMRFamide
 - in *Aplysia*, 84
 - in buccal ganglia of *Aplysia*, 37, 67
 - as Ca²⁺ modulator, 61
 - and cell-adhesion molecules, 62
 - as heterosynaptic modulator, 60
- Force generation
 - biological mechanisms for, 142
 - by hydrostatic skeletons, 143–144
 - by muscular hydrostats, 144–145
 - by musculoskeletal systems, 145–146
- Frogs, 147
- Generalization, 27
- Getting, P.A., 140
- Glutamate
 - in crayfish, 75
 - at NMDA receptor synapses, 18, 90
 - receptors to, 92. *See also* NMDA receptor synapses
- Goldfish, Mauthner cells of, 97
- Graded synaptic transmission, 126, 140
 - and hexapod locomotion, 150
- Gradient descent, 38, 111, 167
- Guinea pig, 81
- Habituation, 61
- Hebb, D.O., 85
- Hebbian learning rules and synapses
 - activity signal of, 68
 - and activity trace, 89
 - definition of, 85

- Hebbian learning rules and synapses (*cont.*)
 - formalisms describing, 86–89
 - in hippocampus, 86
 - and LTP, 69, 86
 - mechanisms for, 89–92
 - retrosynaptic transmission at, 41, 68
 - single-locus requirement of, 68
- Heterosynaptic depression, 82, 83
- Heterosynaptic modulation, 60–61
- Heterosynaptic potentiation, 82
- Hexapod locomotion, 149–161
 - and central pattern generation, 150
 - degrees of freedom of, 160
 - distributed pattern generation, 150, 163
 - metachronal wave, 157
 - network controller for, 150–161, 155–159
 - gaits, 157–161, 162
 - sensory input to, 157–161
 - and peripheral plant, 155–159
 - phases of, 156
 - robot using, 160–161
 - tuning for periphery of, 160
 - slower gaits, 151
 - stance reflex, 156
 - tripod gait, 151
- Hidden layers, 25, 166, 172, 175
- Hippocampus. *See also* Long-term potentiation (LTP): in hippocampus
 - activity-dependent neuromodulation in, 98
 - arachidonic acid in, 60
 - CA1 region
 - of, 63, 79, 80, 81, 82, 86, 87, 90, 92, 95
 - CA3 region of, 76, 82
 - dentate gyrus of, 82, 83, 86, 95
 - heterosynaptic depression in, 82
 - models of, 108
 - modification threshold in, 98
 - synaptic strengths in, 21
 - theta waves of, 169
- Histamine, as heterosynaptic modulator, 60
- Horses, 141
- Hydrostatic skeletons, 143–144
 - of *C. elegans*, 143
 - of leeches, 143
- Identification models, 5, 166
- Information processing, 21
 - and information storage, 103
- Input/output function. *See* Neuronal input/output function
- Instrumental conditioning, 104
- Intracortical axons, 16
- Inverse-Hebbian. *See* Anti-Hebbian rules
- IPSPs
 - in buccal ganglia of *Aplysia*, 30, 45
 - in leech, 119
- Joints
 - degrees of freedom of, 146
 - dynamics of, 146
 - kinematics of, 146
- Kandel, E.R., 10
- Learning. *See also* Synaptic plasticity; Synaptic weights: changes to mechanisms of, 71
- Learning by example, 6
- Learning rate, 76, 80, 87
 - negative values for, 80
- Learning rules
 - associative
 - activity-dependent neuromodulation. *See* Activity-dependent neuromodulation
 - BCM rule, 95–98
 - combining potentiation and depression, 87
 - covariance rule, 94–95
 - definitions of, 72
 - examples of, 85–102
 - associative reward-penalty, 189
 - associative vs. nonassociative, 69, 71
 - backpropagation. *See* Backpropagation
 - changing weights by, 27
 - classifications of, 69
 - and forms of learning, 104
 - Hebbian. *See* Hebbian learning rules and synapses
 - monitored vs. unmonitored, 111
 - and neuronal circuits, 104, 140
 - nonassociative, 72–85
 - definitions of, 71
 - heterosynaptic, definition of, 72
 - heterosynaptic, examples of, 82–85
 - heterosynaptic, mechanisms of, 84
 - homosynaptic, definition of, 72
 - homosynaptic, examples of, 72–81
 - postsynaptic activity-dependent, 80–81
 - presynaptic activity-dependent, 72–80
 - postsynaptic competition, 69
 - postsynaptic growth, 69
 - presynaptic competition, 69
 - random search strategies, 189
 - RTRL algorithm, 168

- Leech
 - hydrostatic skeleton of, 143
 - local bending reflex of. *See* Local bending reflex of leech
- Lizards, 142, 144, 149
- Lobster gastric mill CPG, 114
 - oscillations in, 170
- Local bending reflex of leech, 108, 113–114, 172
 - background of, 116–119
 - dedicated interneuron model, 120
 - distributed processing in, 123–124
 - input-output function of, 119
 - interneuron number needed for, 134
 - interneurons, identification of, 120
 - model analysis, 132–134
 - model architecture, 124–125
 - model performance, 130–134
 - multiple interneurons of, 122
 - time-dependent representation of, 132–134
- Locomotion, 141
 - hexapod. *See* Hexapod locomotion
- Logic gates, 5
- Long-term depression
 - in *Aplysia*, 79
 - in cerebellum, 92
 - in cortex, 92
 - in crayfish, 79
 - definition of, 79
 - in hippocampus, 79
 - low-frequency induction of, 79
 - pattern dependence of, 90
 - requirement for Ca^{2+} of, 90, 92
 - voltage dependence of, 95
- Long-term potentiation (LTP)
 - amplitude of, 77
 - in *Aplysia*, 76
 - CO as retrosynaptic transmitter for, 63
 - in crayfish, 75, 79
 - definition of, 62
 - and dendritic spines, 92
 - duration of, 75
 - and Hebbian synapses, 69, 86
 - in hippocampus, 62, 76, 86
 - to increase gain, 79
 - NMDA receptors in, 63
 - NO as retrosynaptic transmitter for, 63
 - nonassociative, 75–79
 - optimal stimulus for, 79
 - pattern dependence of, 90
 - prevalence of, 75
 - quantal analysis of, 51
 - range of, 54
- requirement for Ca^{2+} of, 63
- requirement for retrosynaptic transmission of, 63
- in sympathetic ganglia, 76
- voltage dependence of, 95
- Martin, K.A.C., 7
- Massively parallel processors, 5
- Metachronal wave, 157
- Mg^{2+} ions
 - and NMDA receptor synapses, 18–19
 - voltage-dependent block by, 18–19
- Mobilization of transmitter, 101
- Model networks. *See also* Networks
 - compartmental, 108–110
 - single vs. multiple, 126
 - deep vs. shallow, 42
 - hybrid, 116
 - of local bending reflex, 116–134
 - for locomotion, 149–161
- Modification threshold, 96–98
- Modularity, 8
- Multilayer architecture, 2
- Multiple-electrode arrays, 4
- Muscles
 - and joints, 145, 147
 - as nonlinear actuators, 146
 - properties of, 142–143
- Muscular hydrostats, 144–145
 - force conversion by, 145
- Musculoskeletal systems, 145–146
- Myomodulin
 - in buccal ganglia of *Aplysia*, 37
 - as heterosynaptic modulator, 61
- Nervous system
 - function of, 1
 - interaction with periphery of, 161–162
- Networks
 - analysis of, 3
 - anatomy vs. function of, 6
 - architecture of, 25. *See also herein* multilayer architecture of
 - independence from biophysics of, 4
 - beyond neurobiology, 135
 - and brain algorithms, 13, 107
 - of buccal ganglia of *Aplysia*, 32–33
 - class one, 17
 - class two, 17–19, 33
 - classification by synapses of, 17–19
 - components of, 23
 - computation vs. construction of, 6
 - and distributed processing, 107

- Networks (*cont.*)
- dynamic recurrent. *See* Dynamic recurrent networks
 - dynamic vs. static synapses of, 17–19
 - dynamics of, 141
 - elements of, 23. *See also* Units
 - emergent properties of, 163
 - feedback in, 140. *See also* Dynamic recurrent networks; Recurrent networks
 - first generation of, 1, 107, 138
 - formalism describing, 24
 - generalization by, 177
 - generations of, 1–2
 - genetic algorithms for, 163
 - hybrid, 3, 116, 135
 - of local bending reflex, 116–134
 - instantiations of, 29, 175
 - and biological synaptic strengths, 33
 - of leech, 113–114
 - modification by experience, 71, 163
 - multilayer architecture
 - of, 2, 10, 25, 110, 138
 - neurobiological equivalents of, 25, 32
 - neuromorphic, 6
 - neurons of, 23. *See also* Units
 - nonlinear dynamics of, 141
 - optimization algorithms, 107, 110. *See also* Learning rules
 - oscillating, 169–172
 - and peripheral plants, 141–142, 156–161
 - polymorphic, 140
 - recurrent. *See* Dynamic recurrent networks; Recurrent networks
 - reviews of, 22
 - second generation of, 2, 10, 138
 - silicon implementation of, 11
 - static vs. dynamic synapses of, 17–19, 140, 188
 - synaptic weights of. *See also* Synaptic weights
 - importance of, 27
 - non-uniqueness of, 29, 175
 - third generation of, 2–5
 - advantages for neurobiology of, 3–4
 - capabilities of, 4–5
 - circuitry of, 10
 - training of, 17
 - training vs. test phases of, 10
 - VLSI implementation of, 11
 - weights of. *See* Synaptic weights
- Neural integration and differentiation, 184–187
- Neural networks. *See* Networks
- Neural throng, 140
- Neurobiological methods
- ablation, 114
 - dye injection, 122
 - ensemble averaging, 53
 - evoked-potential recording, 4
 - field-potential recordings, 4
 - multiple-electrode arrays, 4
 - and network models, 189
 - optical imaging, 4, 31
 - quantal analysis, 51–58
 - semi-intact preparation, 117, 148
 - simultaneous intracellular recordings, 43, 116
 - single-unit recording, 3, 165, 189
 - synaptic pairing, 45
 - tissue culture, 4
 - variable artificial conductance, 4
 - voltage clamping, 21, 30, 45
- Neuromodulation, 84. *See also* Activity-dependent neuromodulation
- in buccal ganglia of *Aplysia*, 37
 - dynamic reconfiguration by, 139
 - and modification threshold, 97
 - in stomatogastric ganglion, 148
 - in visual cortex, 97
- Neuromorphism
- capabilities of networks with, 5
 - dynamic synapse requirement for, 19
 - of network models, 2
 - scaling requirement for, 19
 - and sign change of weights, 32
 - targeting solution space by, 6
- Neuromuscular junction, 95
- Neuronal density of cortex, 14–17
- Neuronal input/output function, 24, 26, 72, 110, 128, 139, 167, 189
- Neuronal interconnectedness
- independence of cortical size, 16
 - theory of, 14–17
- Neurons
- Betz cells, 16
 - channels of, 139
 - compartmental models of, 108–110
 - single vs. multiple, 126
 - complex cells of visual cortex, 113
 - identified, 22, 30, 116–117, 148
 - individual, 6
 - intrinsic properties of, 139
 - Mauthner cells, 97
 - in networks, 1, 138–141
 - as nonlinear dynamic systems, 139
 - pacemaker, 150–152, 161
 - premotoneuronal, 173

- pyramidal cells, 108
- spinal motor neurons, 16, 148
- thalamic, 139
- Nitric oxide
 - and activity-dependent neuromodulation, 69
 - as retrosynaptic transmitter, 63
- Nitric oxide synthetase, 63
- NMDA receptor synapses, 6
 - as class two, 18–19
 - in LTP, 63, 90
 - mechanism of, 18–19, 90
 - Mg²⁺ block of, 18–19, 90
 - requirement for Ca²⁺ of, 90–92
 - and theta rhythm, 18
 - in visual cortex, 96
 - voltage dependence of, 18
- NO. *See* Nitric oxide
- Nonlinear dynamics, 71, 139
- Nonparametric problems, 6

- Octopus, 144
- O'Dell, T.J., 10
- Olfactory bulb, models of, 108
- Olfactory system, 17
- Operant conditioning, 104
- Optical imaging, 4

- Parietal cortex, spatial processing
 - in, 112, 166
- Patterned motor behavior
 - determinants of, 137–138
 - examples of, 137, 169
 - flexibility of, 137
 - from neural-peripheral interactions, 161
- Pavlovian conditioning, 104
- Perceptron, 2, 138
- Peripheral plant
 - complexity of, 138, 141–142
 - constraints of, 141
 - definition of, 138
 - for hexapod locomotion, 155
 - and nervous systems, 148–149, 155–159
 - and neural networks, 141–142, 148–149
 - nonlinearity of, 138
 - versatility of, 141–142
- Periplaneta americana*, 155. *See also* Cockroaches
- Plasticity
 - extrasynaptic, 6, 71
 - and kinases, 84
 - of preexisting circuitry, 140
 - and second-messenger systems, 84
 - synaptic, 71. *See also* Synaptic plasticity
- Plateau potentials, 148
- Pleurobranchaea sp.*, 148
- Postsynaptic specification of presynaptic release, 50–64
- Posttetanic potentiation, review of, 75
- Primates. *See also* Squirrel monkey
 - brain size of, 13
 - step-target tracking in, 173–180
- Protein kinase A, 84
 - in activity-dependent neuromodulation, 101
- Protein kinase C, 84
- Pyramidal cells, 108

- Quantal analysis, 51–58
 - assumptions of, 51
 - of crayfish LTP, 76
 - direct vs. indirect, 51
 - of hippocampal LTP, 51, 63
 - and membrane noise, 54
 - parameter definitions for, 52–54
 - sources of artifact, 54
 - sources of variance, 56
 - and stationarity, 53, 54
- Quantal content, 51

- Rats, 79
- Recurrent backpropagation. *See* Backpropagation: recurrent
- Recurrent networks, 112
 - connectivity of, 166
 - genetic algorithms for, 163
 - with local feedback, 32
 - modularity in, 8
 - for sample-and-hold function, 181
 - subgrouping in, 8
- Retina, 113, 119
- Retrosynaptic transmission
 - backpropagation and requirement for, 38–42, 65–68
 - in buccal ganglia of *Aplysia*, 58–70
 - by cell-adhesion molecules, 61
 - at Hebbian synapses, 68
 - by heterosynaptic modulation, 60–61
 - morphological mechanisms for, 61
 - in network models, 64–70
 - neurobiological mechanisms for, 58–70
 - postulated existence of, 23
 - purpose of, 40
 - by reverse neurotransmitters, 60
 - in stomatogastric ganglion, 64
 - trophic role of, 64
 - in *Xenopus*, 64
- Reverse-Hebbian. *See* Anti-Hebbian rules

Robotics, 145
 Robots
 deformable actuators for, 145, 163
 hexapod, 160–161
 neural network controller for, 160–161
 tuning for periphery of, 160
 Rodents, brain size of, 13
 Rumelhart, D.E., 42

 Saturation, 18
 Scaling function. *See* Neuronal input/output function
 Scaling of cortex, 13–17
 SCP (small cardioactive peptides)
 in buccal ganglia of *Aplysia*, 37
 as heterosynaptic modulator, 60
 Self-inhibition, 30, 32
 dual receptors for, 60
 Sensitization, 61
 Serotonin
 in buccal ganglia of *Aplysia*, 37, 67
 as Ca^{2+} modulator, 61
 as heterosynaptic modulator, 60, 84, 99, 102
 increased transmitter release by, 61
 Shape from shading, 113, 165
 Sherrington, C.S., 148
 Short-term memory, 169, 180–184
 Shrews, brain size of, 13
 Sialic acid, 62
 Sigmoid function. *See* Neuronal input/output function
 Single-unit recording, 3
 Solution space, 6
 Specification of synaptic strength, 45–48
 Speech recognition, 169
 Spike broadening, 67
 Spinal cord
 circuitry of, 31
 synaptic strengths in, 21
 Squashing function. *See* Neuronal input/output function
 Squirrel monkey, brain size of, 14
 Stance reflex, 156
 Stellate ganglion, 82
 Stent, G.S., 94–95
 Step-target tracking, 169, 173–180
 Stochastic processes, 10
 Stomatogastric ganglion of crab, 64, 148
 Subgrouping, 8
 Symmetric weights of Boltzmann machine, 41
 Sympathetic ganglia, 81
 Synapses. *See also* Electrotonic synapses
 complexity of, 139
 conventional as class one, 18
 dynamic vs. static, 17–19
 electrotonic, 36
 models of, 126–128
 NMDA as class two, 18–19
 static vs. dynamic, 17–19
 Synaptic conductances
 and synaptic weights, 128
 time integral of, 33
 Synaptic currents
 limits on, 128
 and synaptic weights, 17
 Synaptic plasticity. *See* Neuromodulation
 activity-dependent neuromodulation, 65
 associative and nonassociative, 31, 71
 Ca^{2+} –dependence of, 61
 and cell-adhesion molecules, 62
 and excitability, 67
 firing pattern, dependence on, 77
 Hebbian mechanisms for, 89–92
 and kinases, 84
 and learning, 71
 long-term, 62. *See also* Long-term depression; Long-term potentiation (LTP)
 and network design, 163
 and second-messenger systems, 84
 selective induction of, 102
 stimulus specificity of, 104
 and transmitter mobilization, 67, 102
 Synaptic strength. *See* Biological synaptic strength; Synaptic weights
 Synaptic weights
 biological presetting of, 140
 and biological synaptic strength, 7, 33
 changes to. *See also* Learning rules
 activity-dependent, 72–81
 associative, 85–102
 by backpropagation, 40, 167
 covariance rule for, 94–95
 by error signal, 40, 167
 independence of, 49
 by learning rules, 27, 110
 manual adjustment for, 27, 110
 by neuromodulators, 84
 nonassociative, 72–85
 by retrosynaptic messengers, 65
 by RTRL algorithm, 168
 training procedure for, 27
 definition of, 2
 importance of, 7, 27
 of local bending reflex model, 129–130
 negative, avoidance of, 80

- non-uniqueness of, 29, 175
- saturation of, 18
 - avoiding, 77, 87
- scaling inputs by, 23, 167
- sign constrained, 183
- of step-target tracking model, 175–180
- and synaptic conductances, 34
- and synaptic currents, 17, 128

Teacherless algorithms, 41

Thalamus, 139

Third generation networks, 2–5

Time constant, membrane, 108, 128

Tissue culture, 4

Top-down analysis, 7, 165

Trace, activity. *See* Activity trace

Transmitter mobilization, 67

Tritonia sp., 140, 148

Tupinambis nigropunctatis, 144, 149

Turing machines, 169

Turtles, 141

Units

- binary, 110

- formalism defining, 24, 110, 167

- layers of, 25, 110

- neurobiological description of, 23

- sigma-pi, 8, 110

- sum-of-products, 8, 110

- threshold of, 167

V1. *See* Visual cortex: primary

Variable artificial conductance, 4

Varicosities, 61

Vestibulo-ocular reflex,

- 113, 114, 169, 184–187

Visual cortex

- complex cells of, 113

- primary, 14–17

- synaptic depression in, 95

Voltage-dependence of Mg^{2+} block, 18–19

VOR. *See* Vestibulo-ocular reflex

Whales, brain size of, 13

Xenopus, 64

XOR. *See* Exclusive-or

Zipser, D., 42